Tracking with MIMO sonar systems: applications to harbour surveillance

Tracking with MIMO sonar systems: applications to harbour surveillance

Journal:	*IET Radar, Sonar & Navigation*
Manuscript ID	RSN-2016-0080.R1
Manuscript Type:	Research Paper
Date Submitted by the Author:	n/a
Complete List of Authors:	Pailhas, Yan; Heriot Watt University, School of E.P.S. Houssineau, Jeremie; Heriot Watt University, School of E.P.S. Petillot, Yvan; Heriot Watt University, School of EPS Clark, Daniel; Heriot Watt University, Institute of Biological Chemistry, Biophysics & Bioengineering
Keyword:	SONAR, MIMO SYSTEMS, TARGET TRACKING, SONAR SIGNAL PROCESSING
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
MIMO_v02.3.tex	

Tracking with MIMO sonar systems:

applications to harbour surveillance

Pailhas, Y.[1,*], Houssineau, J.[1], Petillot, Y. R.[1], Clark, D. E.[1]

[1]School of Engineering & Physical Sciences, Heriot-Watt University, Edinburgh, UK
[*]Y.Pailhas@hw.ac.uk

Abstract: MIMO (Multiple-Input Multiple-Output) sonar systems offer new perspectives for area surveillance especially in complex environments where strong multi-path and dense clutter can become very challenging. This paper proposes a MIMO sonar system based scheme to tackle the difficult problem of harbour surveillance. An emphasis is put on recognition and tracking on low profile mid-water targets. First, a MIMO simulator which can compute synthetic raw data for any transmitter/receiver pair in multipath and cluttered environment is presented. We show that moving targets such as boats or AUVs (Autonomous Underwater Vehicle) can also be introduced into the simulation environment. The paper then proposes two radically different methods for the underwater target tracking problem in complex environment: a digital tracker and an analog tracker. On the digital side, an implementation of the recently developed HISP filter (Hypothesised filter for Independent Stochastic Populations) is presented. This filter enables robust multi-object tracking as well as track classification capabilities without the use of heuristics. An analog filter based on acoustical time reversal techniques is also introduced. This filter directly uses the returned acoustic field from the scene to focus the sound on the expected target position, hence improving the SNR on the target in complex environments and taking full advantage of the MIMO architecture. The proposed modified time reversal technique is shown to match the prediction / data update steps of a traditional tracking filter.

1. Introduction

Military and civilian harbour protection is one of the priority for defence against terrorism [1]. Presently, harbour surveillance is mainly performed using radar systems, but the heterogenous potential threats, including divers and/or AUVs (Autonomous Underwater Vehicles), requires underwater surveillance as well as a surface surveillance. The recent NATO trials on harbour protection conducted by CMRE (Centre for Maritime Research and Experimentation), such as the HPT06 trials [2], emphasise the need for solutions to this problem. From an acoustical point of view, harbours are particularly challenging to survey for two main reasons: harbours tend to be heavily cluttered by dumped man-made objects and harbours are by definition a very shallow water environment. Those two factors proved to be problematic in an automatic target detection framework. From an operational point of view, the difficulties rise from (1) the variability of the potential threats (e.g. boat, speed boat, AUVs, divers) mixed with mistaking innocent intrusions (e.g. sport divers, tourists), and (2) the large number of vessels entering and exiting this busy thoroughfare at

every time of the day.

If harbour surveillance radars are the primary surveillance tool, the harbour security can also rely on the Automatic Identification System (AIS) to monitor the large commercial vessels entering and exiting the harbour. Several diver detection sonar systems are already available commercially. However, diver detection and tracking is still an active area of research for harbour surveillance and diver rescue applications. Two main approaches have been investigated so far focusing on active sonars or passive arrays. Most active solutions focus on mono-static forward looking sonar [3–5] and image processing approaches. Detection with passive systems use prior knowledge on the diver acoustic behaviour and in particular the periodic sound produced by the diver breathing apparatus [6–10]. However, to the authors' knowledge, no systems have been designed to tackle both the multiplicity of underwater threats and the multi-target problem inherent to a complex harbour scenario. This paper investigates the potential for MIMO (Multiple-Input Multiple-Output) sonar systems to secure an harbour environment.

In recent years, MIMO sonar systems have raised a lot of interest, especially in the anti-submarine warfare (ASW) community. Multistatic sonars outperform monostatic sonar systems in target localisation and detection performances [11, 12]. CMRE, for example, developed a deployable low frequency SIMO (Single Input Multiple Outputs) sonar system called DEMUS. This system has shown enhancements compared to classical methods in detection and tracking of large underwater targets [13, 14]. The few multi-static systems proposed recently however have a very low numbers of sensors and focus more on open water problems rather than small enclosed areas such as harbours. Recently, the authors in [15] exploited the statistical principles inherent to large MIMO sonar systems with spatially distributed antennas and showed increased capabilities in terms of ATR (automatic target recognition) and imaging compared with traditional systems. The conclusions to this new MIMO paradigm lead naturally to consider such systems for surveillance purposes [15–17]. The present paper focuses specifically on the capabilities of spatially distributed MIMO sonar systems for harbour surveillance from a design and an algorithmic point of view.

Harbours are by definition a cluttered and busy environment (presence of lot of boats, fish, school of fish, etc). It is then likely that the MIMO images will be equivalently cluttered and will provided numerous detections and false alarms. The number of detection may prevents the operators from understanding the overall situation directly; robust and efficient tracking algorithms have to be used to provide more accessible representations of the scene. Two classes of multi-object tracking algorithms can be identified within the literature. The first class is made of multi-object filters that are more or less direct generalisations of the Kalman filter and often rely on heuristics to handle creation and deletion of tracks. The advantage with these filters is that they provide track identities, which makes useful information available such as objects' trajectory. The MHT (Multiple Hypothesis Tracker) [18] is one of these filters and has been predominantly used in the field of multistatic sensors [13, 14]. The second class of multi-object filters has been more recently introduced and consists of methods that are based on spatial-statistics tools such as point processes. These tools enable a principled estimation for appearing and disappearing objects at the cost of the loss of track identities. Examples of multi-object filters that belong to this second class are the Probability Hypothesis Density (PHD) filter [19] and the Cardinalised PHD (CPHD) filter [20]. In parallel with these multi-object tracking techniques, it is worth mentioning the use of neural networks. Although neural networks work extremely well in problems such as object localisation and classification [21–23], they can also be useful in a multi-object tracking scheme at the detection stage [24] or the data association problem [25, 26].

In order to overcome the heuristics introduced by the MHT and the loss of track identities

inherent to the PHD, a novel way of representing multi-object systems has been introduced in [27]. One of the filters that has been derived from this novel representation is called the HISP filter [27, 28]. This filter performs joint detection and tracking, to which classification can be added naturally. Indeed, for the same reasons as the HISP filter can distinguish objects, it can also perform classification by distinguishing between different behaviours. The classification capabilities of the HISP filter have already been studied for different applications [16, 29]. This paper presents an implementation of the HISP for simultaneously track and classify different class of targets in a heavily cluttered environment using a MIMO sonar system.

Tracking in the observation space provides extremely encouraging results in both tracking and recognition as it will be shown in section 4.3. While the HISP digital tracker benefits from the high resolution images and as a consequence the high detection rate of a large MIMO sonar system, it is also restricted by its input and does not benefit from the full flexibility of such system. One can argue that the tracking is performed on the resulting MIMO image alone, and therefore is tied to the imagery process. Three main limitations then appear: (1) the tracking process uses only detection, (2) the sensing is not adaptive to the environment nor the target and (3) finally the imaging process assumes a propagation model. The multiplicity of sensors of a large MIMO sonar system however brings as many degrees of freedom as MIMO pairs [15]. To overcome the limitations described above, the present paper proposes a tracking algorithm based on exploiting some of these degrees of freedom using Time Reversal techniques. First developed by Fink [30], Time Reversal operators can be used to focus the energy on a static target [31] independently of any knowledge of the environment. A series of at-sea experiments have demonstrated the feasibility of time reversal approaches despite the fluctuations of the medium [32, 33] . In particular, in a time reversal experiment conducted with the CMRE in 2003, Roux in [34] shows that the coherence was maintained over several hours using the SA-RA system at 3.5kHz. This paper proposed a modified version of Time Reversal based on defocusing to track a moving target. The tracking is then performed at the sensor level, independently from any imagery process.

The objective in this article is to show the surveillance capabilities provided by MIMO sonar systems when combined with appropriate detection, tracking and classification techniques. For this purpose, a realistic MIMO-sonar simulator is introduced and the performance of the HISP filter is assessed on data generated by the simulator. An alternative approach yielding an enhanced signal-to-noise ratio is also presented.

The paper is organised as follows: Section 2 presents a full 3D MIMO simulator which can compute synthetic raw signals for any transmitter/receiver pair configuration in a multipath and cluttered environment. Computing raw echoes in a 3D environment rises a number of computational hurdles. Section 2.4 describes some of the techniques that made the MIMO simulator computationally efficient. In Section 3, an hostile harbour intrusion scenario is presented. Two original solutions for the detection and the tracking of underwater targets using a MIMO architecture are then proposed:

1. The high resolution MIMO images can provide reliable detections. These detections feed into a digital tracker, the HISP filter, which not only track but also classify automatically the multiple targets (Section 4.3).

2. The MIMO sonar system can focus the sound on the moving targets only and track them via a modified version of Time Reversal, increasing signal-to-noise ratio and therefore the probability of detection (Section 5).

Section 4 presents an overview of the HISP filter as well as the inherent principles enabling the

classification using this filter. Finally, Section 5 shows that large MIMO sonar systems offer an ideal platform for time reversal techniques. In particular an unfocussed time reversal mirror algorithm capable of tracking automatically moving targets is developed.

2. MIMO simulator

This section extends the work on simulated MIMO sonar system done in [16]. It describes the three main blocks of the MIMO simulator: the seabed bathymetry generator, the computation of the bistatic reverberation level and the wave propagation model.

2.1. Simulator structure

The MIMO simulator is organised in two main parts. At first, a fractal seabed terrain is generated as explained in section 2.2. Each seabed cell is then randomly populated by scatterer points. The echo from the environment (including seabed reverberation, targets and multipath) is then computed for each transmitter/receiver MIMO pair. The output of the simulator is the full MIMO response in the time domain. The details regarding the acoustic wave propagation computation can be found in section 2.4.

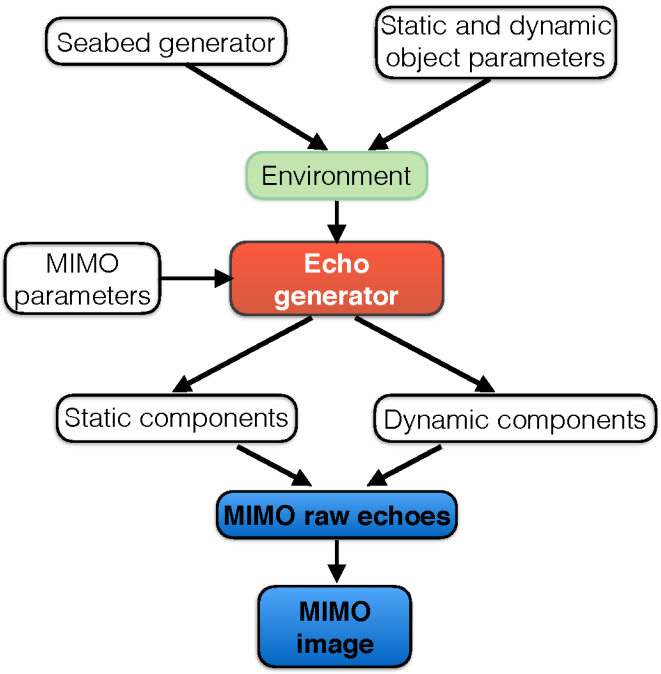


Fig. 1: (Colour online) General structure of the MIMO simulator.

The overall structure of the simulator is pictured in Figure 1. The output of the seabed generator and the various object parameters (structure, dynamic,etc) form the full synthetic environment. The synthetic environment and the MIMO parameters (such as the transducer locations) feed the echo generator module (highlighted in red in the figure). Note that the acoustic wave propagation module (echo generator) can compute independently the static component and the dynamic component of the echo. The reason behind this choice is that the static component has only to be computed once per scenario. The simulator can provide two different outputs (highlighted in blue):

the MIMO raw echoes or the resulting MIMO image (section 2.5).

2.2. Seabed interface

Natural surface geometries are well represented by fractal models [35]. Hence the seabed elevation is modelled through a 2D fractional Brownian motion (fBm) using the Incremental Fourier Synthesis Method developed by Kaplan and Kuo in [36]. The main idea is to model the 1st and 2nd order increments I_x, I_y and I_2 given by the following equations:

$$I_x(m_x, m_y) = B(m_x + 1, m_y) - B(m_x, m_y) \quad (1)$$

$$I_y(m_x, m_y) = B(m_x, m_y + 1) - B(m_x, m_y) \quad (2)$$

$$I_2(m_x, m_y) = B(m_x + 1, m_y + 1) + B(m_x, m_y) - B(m_x, m_y + 1) - B(m_x, m_y + 1) \quad (3)$$

where B is the 2D fBm and (m_x, m_y) represent the pixel location. Those 1st and 2nd order increments can be computed thanks to the FFTs (Fast Fourier Transforms) of their correlation functions. The 1st and 2nd order increment FFTs correlation functions are given by:

$$S_2(\omega_x, \omega_y) = \frac{32\sqrt{\pi}\sin^2(\omega_x/2)\sin^2(\omega_y/2)\Gamma(2H+1)\sin(\pi H)}{\sqrt{\omega_x^2+\omega_y^2}^{2H+2}} \quad (4)$$

$$S_x(\omega_x, \omega_y) = \frac{S_2(\omega_x,\omega_y)}{4\sin^2(\omega_y/2)} \quad (5)$$

$$S_y(\omega_x, \omega_y) = \frac{S_2(\omega_x,\omega_y)}{4\sin^2(\omega_x/2)} \quad (6)$$

where H is the Hurst parameter and is directly related to the fractal dimension of the fBm. (ω_x, ω_y) are the spatial frequencies. The fBm is then obtained by adding up the increments given by Eq. (1), (2) and (3). Figure 2 displays an example of 2D fractional Brownian surface generated using this technique.

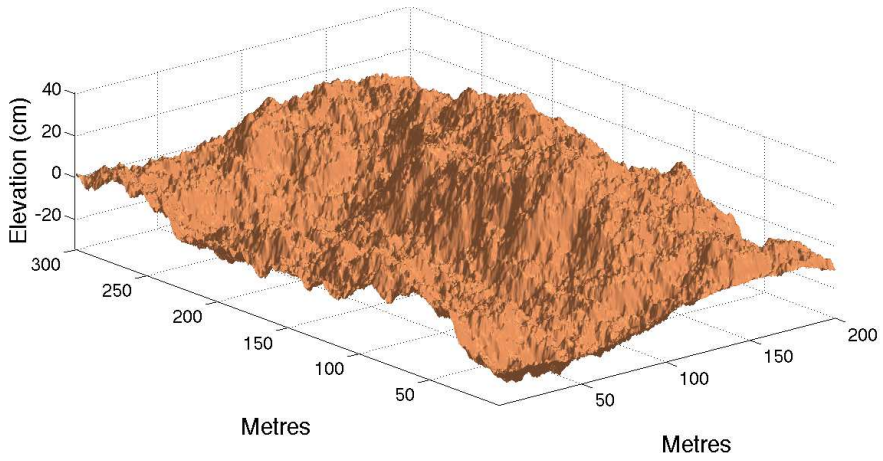


Fig. 2: (Colour online) Example of 2D fBm with H = 0.8 (fractal dimension = 2.2)

2.3. Bistatic reverberation level

The bistatic scattering strength S_b is computed using the model developed by Williams and Jackson from APL-UW [37]:

$$S_b(\theta_s, \phi_s, \theta_i) = 10\log[\sigma_{br}(\theta_s, \phi_s, \theta_i) + \sigma_{bv}(\theta_s, \phi_s, \theta_i)] \quad (7)$$

where $\sigma_{br} = [\sigma_{kr}^{\eta} + \sigma_{pr}^{\eta}]^{1/\eta}$ is the bistatic roughness scattering which includes the Kirchhoff approximation σ_{kr} and the perturbation approximation σ_{pr}. The last term σ_{bv} is the sediment bistatic volume scattering. θ_i and θ_s are respectively the incident and scattered grazing angles, while ϕ_s is the azimuth difference between the incident and scattered fields. S_b then depends on the bistatic geometry of the transmitter/receiver pair as well as the sediment physical properties. Figure 3 displays an example of bistatic scattering strength for a transmitter/receiver pair situated 141m apart and both at 7.5m above the seafloor. In this figure, S_b is computed for two different sediment types (coarse sand and sandy mud) with the same fBm interface. There is around 10dB difference in the bistatic reverberation level between the two sediment types which can play a role in the detection/tracking process [16]. These two sediment types will be considered thereafter for the numerical simulations.

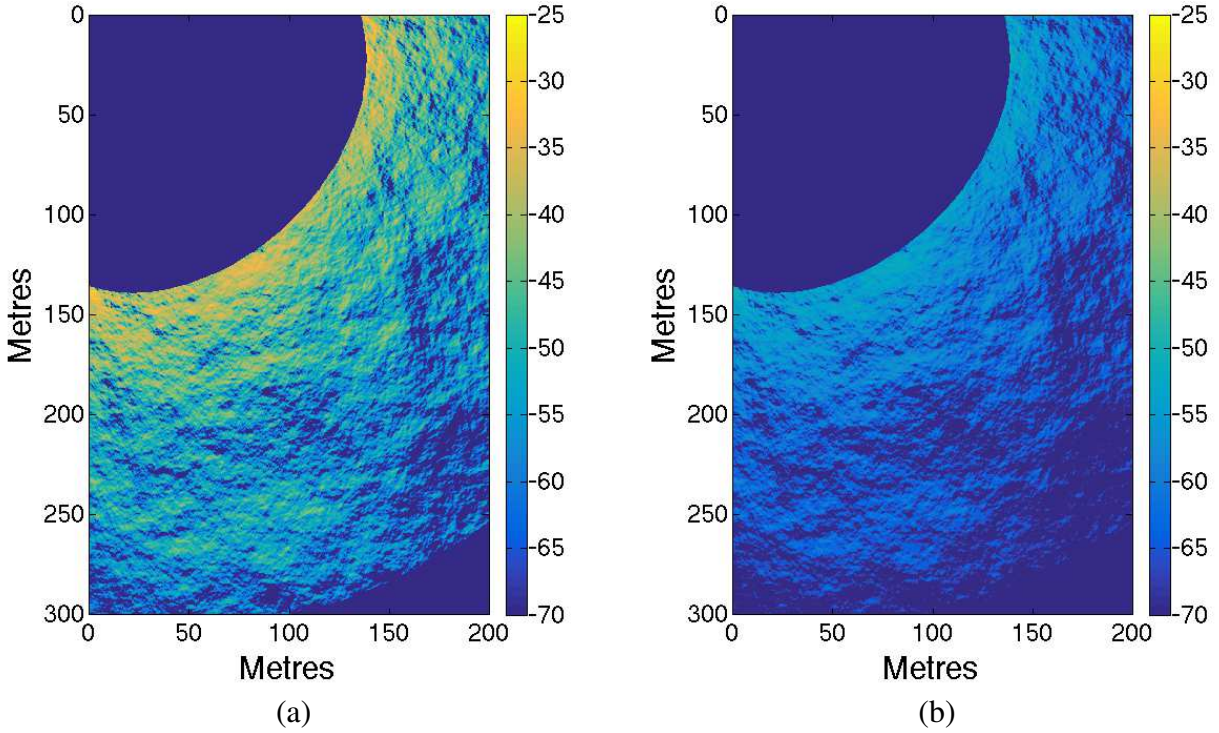


Fig. 3: (Colour online) Bistatic scattering strength in dB relative to one Tx located at [x=0m, y=100m, z=7.5m] and a Rx located at [x=100m, y=0m, z=7.5m] for (a) a coarse sand sediment type and (b) a sandy mud sediment type.

2.4. Acoustic field propagation

This paper focuses on harbour environment. By nature such environments are very shallow water and then subject to heavy multipath. Sound propagation can become extremely complex and problematic. Due to the very shallow water hypothesis, a constant sound speed throughout the environment is assumed. Thanks to the relative short distances involved, the MIMO sonar system considered here will be a high frequency system (above 1kHz). Note that the MIMO sonar system considered in section 3 has a central frequency of 30kHz. At these frequencies, the wave equation can be well approximated by the Eikonal equation [38, 39] and solution to the sound propagation

can be found using Ray tracing techniques [40]. With the constant sound speed assumption, each ray propagates in straight line simplifying greatly the echoes generation and cutting down drastically the computational cost. The same assumption also allows the use of the source mirror theorem for the multipath computation [41]: the multiple trajectories from one source point to a receiver point can be computed geometrically by considering the straight path between the source images obtained by successive bottom and surface reflections and the receiver point, the multiple paths are obtained by *folding* back the multipath images. Figure 4 draws the principles of the Mirror theorem of the geometrically computed multipath: The multipath between the transmitter (black circle) and the receiver (red triangle) are the folded straight lines between the transmitter images and the receiver. The simulator computes a maximum of three bounces. So for a mid-water target, a total of nine multiple paths are computed. The choice behind the limited number of bounces comes from the coherent processing of the MIMO responses described in Section 2.5. As the echoes are added coherently, incoherent echoes resulting from multipath are greatly suppressed [42].

For the same reason, only the coherent component of the sea surface forward scattering is taken into account in the multipath computation. The coherent forward reflection coefficient [37] is approximated by $\exp\left(\frac{4\pi h \sin\theta}{\lambda}\right)$ where h is the surface wave height, θ the reflection angle and λ the wavelength. Note that the Doppler shift due to the wave motion has not been taken into account. The two main reasons for this choice are (1) harbours are usually protected by breakwaters or moles and the average wave height is considerably attenuated compared to the open sea and (2) Doppler tolerant waveforms are used by the MIMO sonar system.

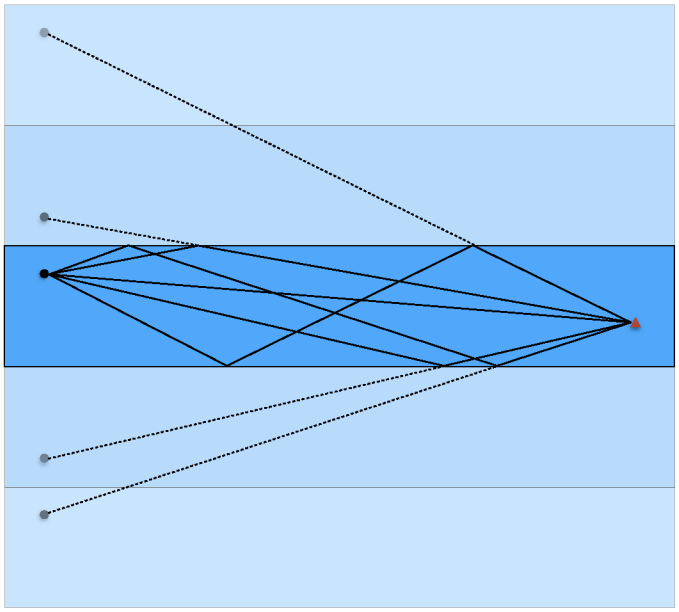


Fig. 4: (Colour online) Multipath computed geometrically using the source images.

To synthesise time echoes, a random scatterer point cloud including random positions and random intensities is generated for each cell in the seabed. The direct bistatic echo $r_{lk}(t)$ of one cell from the transmitter l to the receiver k can be written as follows:

$$r_{lk}(t) = \sigma_b \alpha_{lk} \sum_{q=1}^{Q} \gamma_{lq} \zeta_q s_l(t - \tau_{lq} - \tau_{kq}) \tag{8}$$

where s_l is the pulse send by the transmitter l, τ_{lq} and τ_{kq} are respectively the travelling time from the transmitter l to the scatterer q and he travelling time from the scatterer q and the receiver k, Q is the number of scatterers contained in the cell, ζ_q is the reflecting coefficient of the scatterer q. Note that the overall reflected energy of each cell is normalised: $\zeta_q > 0$ and $\sum_{q=1}^{Q} \zeta_q = 1$. γ_{lq} corresponds to the energy associated with the ray traveling from the transmitter l to the scatterer q, α_{lk} corresponds to the transmission loss along the path transmitter $\rightarrow$ cell $\rightarrow$ receiver, finally σ_b correspond to the bistatic reverberation level which is a function of the relative positions of the triplet (transmitter, receiver, cell) and the normal vector of the cell.

Considering a patch of seabed of 250m $\times$ 250m to be simulated with a cell resolution of 50cm $\times$ 50cm, an average of 20 scatterers per cell and a MIMO sonar system with 10 transmitters and 10 receivers (100 MIMO pairs), the MIMO simulator has to generate around half a billion rays to compute the full MIMO response of this area. Brute force computation using MATLAB on a standard laptop requires around 2 months of computation. To overcome the computational hurdle, two computational tricks were implemented:

2.4.1. Far-field assumption: The first trick comes from the observation that the simulator propagates rays from point cloud to point cloud. Assuming the first point cloud S_1 contains Q_1 scatterers and the second S_2 contains Q_2, the total number of rays to be computed between a starting point Tx to the receiver end Rx interacting with S_1 and S_2 is $Q_1 \times Q_2$. Eq. (8) then becomes:

$$r_{lk}(t) \propto \sum_{q_2=1}^{Q_2} \sum_{q_1=1}^{Q_1} \gamma_{lq_1} \zeta_{q_1} \zeta_{q_2} s_l(t - \tau_{lq_1} - \tau_{q_1 q_2} - \tau_{q_2 k}) \tag{9}$$

However the distance between the two sets of scatterers S_1 and S_2 is generally large and a far-field approach can be considered. The acoustic field originating from the Q_1 scatterers of S_1 and arriving at S_2 is approximated by a single source at the centre of S_1. Eq. (9) can be rewritten as:

$$r_{lk}(t) \propto \sum_{q_2=1}^{Q_2} \zeta_{q_2} s_1(t - \tau_{\mathrm{S}_1 q_2} - \tau_{q_2 k}) \tag{10}$$

with

$$s_1(t) \propto \sum_{q_1=1}^{Q_1} \gamma_{lq_1} \zeta_{q_1} s_l \left(t - \tau_{lq_1} - \tau_{q_1 \mathrm{S}_2} + \tau_{\mathrm{S}_1 \mathrm{S}_2}\right) \tag{11}$$

where $\tau_{q_1 \mathrm{S}_2}$ represents the travelling time between the scatterer q_1 and the centre of S_2 and $\tau_{\mathrm{S}_1 \mathrm{S}_2}$ represents the travelling time between the two sets S_1 and S_2. By using the far-field assumption, the total number of rays to be computed is reduced from $Q_1 \times Q_2$ to $Q_1 + Q_2$.

2.4.2. Sparsity: A limitation factor of Eq. (8), from a computational point of view, is the multiple calculations of the delayed signal. The main tool to compute a delayed signal is the well known FFT property:

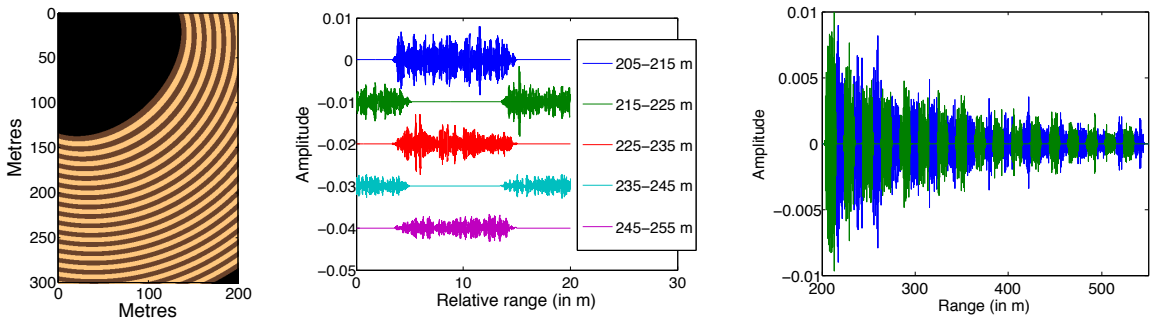
$$f(t - u) \Leftrightarrow \mathrm{e}^{-iu\omega} \hat{f}(\omega) \tag{12}$$

where u corresponds to the delay and $\hat{f}$ the Fourier transform of the function f. Considering the same scenario as above with patch of seabed of 250m $\times$ 250m to simulate, the maximum bistatic

range $r_{\max}$ is around 700m. Considering now the echo related to a single cell (50cm $\times$ 50cm), this bistatic echo is extremely sparse over the potential 700m range signal. The trick here is to use the cell echo sparsity in conjunction with the circular convolution properties of the DFT. Ideally the function f would have to be defined over the time window [0, $r_{\max}/c$] where c is the sound speed in water. Considering now f defined over [0, r_0/c] with $r_0 \ll r_{\max}$, the delayed function $f(t-u)$ becomes

$$f\left(\mathrm{mod}[t-u, r_0/c]\right) \tag{13}$$

Computing Eq. (13) instead of $f(t-u)$ reduces drastically the FFT complexity, the memory needed and, as a result, the computational time. The only ambiguity to solve is the quotient of the $\mathrm{mod}[t-u, r_0/c]$ operation. Figure 5 draws the outlines of the algorithm. To implement efficiently the principle of Eq. (13), the seabed is first divided into range bands. Figure 5(a) shows the environment division into range bands, each colour corresponds to a 10m width range band. The echoes relative to each band are computed independently over a very small window. Figure 5(b) plots the echoes relative to the first five range bands. Note that here $r_0 = 20$m. The range division of the environment not only introduces the sparsity of the signal but also solves the quotient of the modulo operation. The echoes are then recombined to give the full range bistatic response as seen in figure 5(c).



(a) (b) (c)

Fig. 5: (Colour online). (a) Range decomposition of the observed scene in 10m range bands. (b) Individual range band echoes contribution. (c) Full echo response recomposition.

Using the far-field assumption and introducing the sparsity into the echo signal greatly reduce the computational time from two months to few hours [16]. Multiple scenarios can now be synthesised in a reasonable time. Note that the computational time has been estimation using MATLAB on a standard laptop.

2.5. MIMO imaging

The output of the simulator described above is the echo responses of the environment (including seabed reverberation, targets and multipath) for all the MIMO pairs of the system. Different fusion schemes can be implemented depending on the application and the purpose of the MIMO sonar system. In [15] for example, the authors demonstrated the ATR capabilities of large MIMO sonar systems. The target responses were fused using a Bayesian framework to estimate their PDF (probability density function).

For the purpose of this paper, the exact locations of the different transducers of the MIMO sonar system are known and coherent processing of the MIMO echoes is possible. MIMO imaging

can then be achieved using a multistatic back-projection algorithm. Multistatic back-projection algorithm is a variant of the bistatic back-projection algorithm developed by the SAR (Synthetic Aperture Radar) community. Further details can be found in [43]. The traditional back-projection algorithm used in SAS (Synthetic Aperture Sonar) or SAR imaging is computed by integrating the echo signal along a parabola. In the bistatic case the integration is done along ellipses. For the multistatic scenario the continuous integration is replaced by a finite sum in which each term corresponds to one transmitter/receiver pair contribution. Figure 6 displays an example of a MIMO image using the multistatic back-projection algorithm: the seabed is a fractal coarse sand seafloor, a mid-water target (point target) is present at the location [x =200m, y=150m, z=3m].

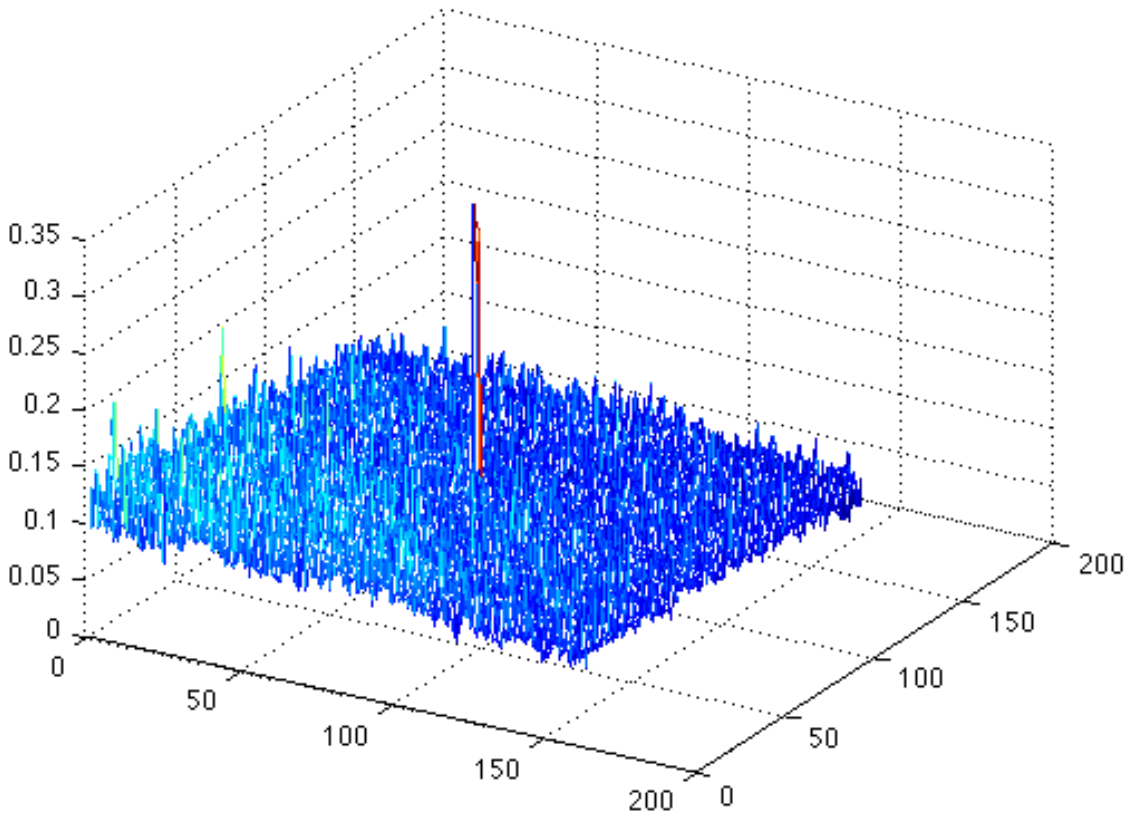


Fig. 6: (Colour online) Surface plot of the synthetic aperture MIMO image of a mid water -30dB target on a coarse sand sediment background [16].

3. Harbour surveillance scenario

The harbour surveillance scenario described in this section has been designed to test the performances of the tracking algorithm presented in Section 4.

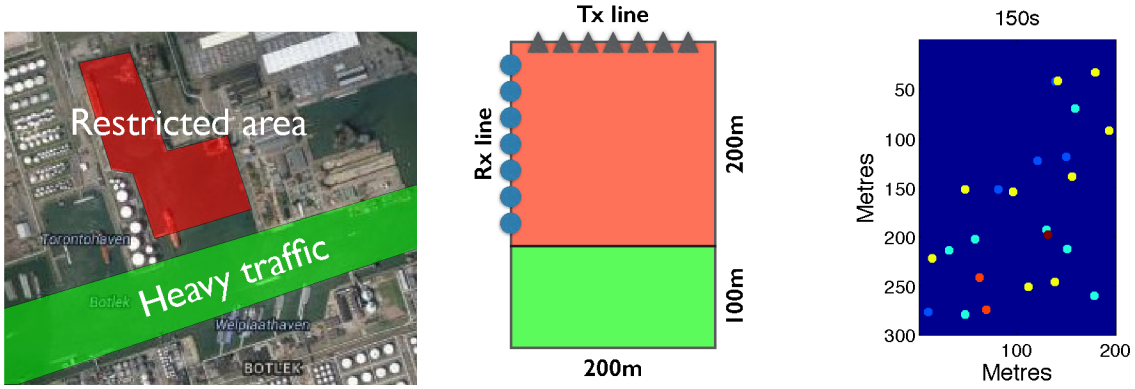
Figure 7a displays the overall scenario: in a harbour environment, a restricted area (highlighted in red in the figure) is located close a traffic area (highlighted in green). The purpose of the MIMO sonar system is to protect the restricted area by detecting any underwater intrusion. The challenge

is to operate in a very shallow and very cluttered environment. A lot of man-made objects usually lie on the bottom of harbours. It is also expected that a number of fish or schools of fish swim freely in the harbour.

Figure 7b displays the geometry of the synthetic environment: 300 $\times$ 200 m area to survey. The colour coded areas match the one introduced in Figure 7a: The red area on the top is the area to protect, and the green area represents the channel where boats can circulate. In Figure 7b, the grey triangles represent the transmitter locations, and the blue disks the receiver locations. The synthetic harbour is in average 15m deep and its bottom is composed of coarse sand or sandy mud sediment. The MIMO sonar systems is composed of 11 transmitters located at the top in the figure and 11 receivers located on the right. The transmitters and receivers are equally distributed with 10m spacing. All the transducers are located at 7.5m depth. The central frequency for the MIMO sonar system is 30kHz and the resolution cell 50cm.

The environment is populated with different targets: fixed man-made objects lying on the bottom, fish, boats and one AUV. The bottom targets are randomly positioned in the environment. The fish swim in the harbour following a Brownian motion. A total of three boats sail in the channel area. One of the boat drops an AUV of whom the mission is to infiltrate the restricted area.

A simple detector based on the reverberation level is performed on the MIMO images. The detection results is feeding the digital tracker presented in section 4. Note that bottom targets and fish are likely to produce detections which can cause heavy clutter. Figure 7c displays an example of input to the multi-object tracker. Note that the detection have been colour coded for display purposes only.



(a) Harbour scenario (b) Geometry of the MIMO simulation (c) Colour-coded detections

Fig. 7: (Colour online) Description of the considered synthetic scenario (7b & 7c) with a real-harbour configuration (7a). Colour code: (light blue) static bottom object, (dark blue) false alarm, (yellow) fish, (orange) boat, (red) AUV.

4. Classification with the HISP filter

The number of detections generated by sonar systems can potentially be large, preventing a direct interpretation from an operator. Suitable algorithms correlating the data in space and time are then used to facilitate the understanding of the scene. The objective of these algorithms is to estimate the number of objects detected by the sensor(s) and to provide information about the state of each of these objects.

Due to the complex nature of the harbour protection problem, robust and reliable estimators are needed. If the tracking filter of choice for multistatic systems has been the MHT [13, 14], because of its ability to propagate track identities through time, it suffers strong reliance on heuristics for the appearance and disappearance of targets and a lack a adaptivity. Even though it has a lower computational complexity, the PHD filter has been shown to outperform the MHT in different situations [44, 45]. An example of the application of the PHD filter methodology to sonar data can be found in [46]. However, there is no direct way of performing classification with the PHD or CPHD filters because of the underlying nature of point processes. In order to obtain principle approaches that naturally maintain track identities, a novel way of representing multi-object systems from which the HISP filter can be derived has been introduced in [27].

In this section, an overview of the principles behind the HISP filter are first introduced, followed by a section on how classification is performed within this filter. Results on the previously introduced harbour surveillance scenario are then presented.

4.1. The HISP filter

The objective in this section is to provide a concise description of the HISP filter. A detailed presentation of the filter can be found in [27, Chapt. 4]. The HISP filter is a Bayesian recursive algorithm performing joint detection and tracking based on point observations of a multi-target system. The filter follows the usual multi-target tracking assumptions, i.e. a) each target's dynamics and observation follow a hidden Markov model, i.e. the dynamics is Markovian and the observation depends only on the current state, b) targets are independent from each other and generate at most one observation per time step following a Bernoulli process, c) the clutter is independent from the targets and d) targets appear anywhere in the field of view and their disappearance follows another Bernoulli process. The HISP filter can be seen as propagating a collection of *hypotheses*

$$\left\{(w_t^{(i)}, p_t^{(i)})\right\}_{i\in\mathbb{I}_t}, \qquad w_t^{(i)} \in [0,1],\ p_t^{(i)} \in \mathbf{P}(\mathbf{X}) \tag{14}$$

consisting of single-object probability laws $p_t^{(i)}$ on the state space $\mathbf{X}$ to which is associated a probability of existence $w_t^{(i)}$, i.e., a probability for a given law to represent a true object. The index set $\mathbb{I}_t$ can be given an explicit expression based on the time of creation and of the observation history of a given hypothesis. The state space $\mathbf{X}$ is defined as the union

$$\mathbf{X} = \mathbf{X}^\bullet \cup \{\psi\}$$

of the subset $\mathbf{X}^\bullet$ of an Euclidean space and of an isolated point ψ. For instance, the set $\mathbf{X}^\bullet$ might be a subset of $\mathbb{R}^4$ describing the position and velocity of the objects in a 2-dimensional scene. The isolated point ψ represents objects which are not yet in the area of interest or which already left it. In consequence, a birth is a transition from ψ to $\mathbf{X}^\bullet$ and the converse is interpreted as a disappearance. The representation (14) specifically describes each possible object as opposed to the Poisson-point-process representation of the PHD filter which models all objects at once. This representation will be useful for addressing the classification problem as explained in Section 4.2.

The time prediction of the HISP filter can be expressed simply as

$$p_t^{(i)}(\mathrm{d}x) \overset{f}{=} \int M_t(y, \mathrm{d}x)\hat{p}_{t-1}^{(i)}(\mathrm{d}y) \tag{15a}$$

$$w_t^{(i)} = \hat{w}_{t-1}^{(i)}, \tag{15b}$$

for all previously existing targets $i \in \hat{\mathbb{I}}_{t-1}$, where M_t is a Markov kernel describing the transition from time $t-1$ to time t and where a hat denotes observation-updated quantities. Note that $\nu(\mathrm{d}x) \overset{\int}{=} \mu(\mathrm{d}x)$ is a shorthand notation for $\int f(x)\nu(\mathrm{d}x) = \int f(x)\mu(\mathrm{d}x)$ for all bounded measurable functions f. Probability laws corresponding to newborn objects are treated separately and a new index set $\mathbb{I}_t$ is introduced as the union of the index sets corresponding to existing and newborn objects. The weight associated with each single-object law is not affected by the time prediction since no information is introduced in this step.

The observation update of the HISP filter has a more involved expression and only a general form of this step is given as a motivation. An observation set Z_t is made available at time t and this set is augmented by the empty observation ϕ in order to represent the case of missed detection. Also, a likelihood function $\ell_z : \mathbf{X} \to [0,1]$ is introduced for each observation $z \in \bar{Z}_t = Z_t \cup \{\phi\}$. For any given $i \in \mathbb{I}_t$ and $z \in \bar{Z}_t$, the update of the hypothesis with index i by the observation z takes the form

$$\hat{p}_t^{(j)}(\mathrm{d}x) \overset{\int}{=} \frac{\ell_z(x)p_t^{(i)}(\mathrm{d}x)}{\int \ell_z(y)p_t^{(i)}(\mathrm{d}y)}, \tag{16}$$

and

$$\hat{w}_t^{(j)} = \frac{w_{\mathrm{ex}}(i,z)w_t^{(i,z)}}{\sum_{z\in\bar{Z}_t} w_{\mathrm{ex}}(i,z)w_t^{(i,z)}}, \tag{17}$$

where j is an index in the observation-updated set $\hat{\mathbb{I}}_t$, where $w_t^{(i,z)} = \int \ell_z(x)p_t^{(i)}(\mathrm{d}x)$ is the *compatibility* between the prior law with index i and the observation z and where $w_{\mathrm{ex}}(i,z)$ is a scalar in the interval $[0,1]$ describing the compatibility between the hypotheses indexed by $\mathbb{I}_t \setminus \{i\}$ and $Z_t \setminus \{z\}$, taking into account false positives and missed detection. One of the advantages with the formulation (17) is that it provides an a posteriori probability of missed detection when $z = \phi$, which enables scenarios with a low probability of detection to be also tackled. Another advantage of the HISP observation update is that it has a better *memory* than the one of the PHD filter: if it is known that there is a single object in the scene before the update and if two of the observations in Z_t are "close" to the assumed position of the object, then the PHD filter might expect up to 2 objects after the update whereas the HISP filter will remember that there was only one object.

Yet, computing the function w_{ex} exactly would be extremely costly and approximations need to be considered. The approximation that is used when computing the term $w_{\mathrm{ex}}(i,z)$ consists in assuming that the law with a given index $k \in \mathbb{I}_t \setminus \{i\}$ is unlikely to be compatible with two different observations $z_1, z_2 \in Z_t \setminus \{z\}$, i.e., it holds that

$$w_t^{(k,z_1)} w_t^{(k,z_2)} \approx 0. \tag{18}$$

In this way, the collection $\{w_t^{(j)}\}_{j\in\hat{\mathbb{I}}_t}$ can be computed with a complexity of the order of $|\mathbb{I}_t||Z_t|$, which is comparable to the complexity of the PHD filter. A detailed expression of the function w_{ex} under this assumption is given in [27, Corollary 4.2].

To handle the increase of the number of hypotheses, pruning and merging procedures can be applied as with the Gaussian-mixture implementation of the PHD filter [47], even though the representation (14) is not a Gaussian mixture. Pruning is performed straightforwardly by removing hypotheses that have a probability of existence below a given threshold τ_{p}. The interpretation of the merging procedure differs from the Gaussian mixture case, since the action of merging probability distributions implies that the two merged hypotheses are then assumed to represent the same

object. This can however be done under the assumption expressed in (18). A merging threshold τ_{m} is also defined for the considered statistical distance.

Track extraction can be performed straightforwardly by picking the single-object laws that have a probability of representing a true object that is above a given threshold τ_{e}. Also, the integrals in (15a) and (16) can be computed either with the Kalman approach or with Monte Carlo methods, providing a Kalman Filter and a Sequential Monte Carlo implementation [28] of the HISP filter.

4.2. Classification

In order to perform classification, distinct behaviours have to be modelled and embedded into the filter. These models are often related to different object dynamics as will be the case in the considered scenario, described in Section 3. For the sake of simplicity, it is assumed that only Markovian dynamics are required to distinguish between objects of interest. This means that the different classes of objects that need to be distinguished must be observable on a single prediction step, i.e., the model that is the best Markovian predictor will define the class of a given object.

More formally, a specific Markov kernel M_t^c is introduced on an appropriate state space $\mathbf{X}_c$ for each class c in the set C of all classes of interest. Also, a collection

$$\left\{(w_{t-1}^{(i,c)}, p_{t-1}^{(i,c)})\right\}_{c\in C}$$

is now associated with each index $i \in \hat{\mathbb{I}}_{t-1}$. The time prediction just consists in applying the Markov kernel from the same class as the considered single-object probability. The only modification of the observation update is in (17) which has to be rewritten as

$$\hat{w}_t^{(j,c)} = \frac{w_{\mathrm{ex}}(i,z) w_t^{(i,c,z)}}{\sum_{z\in\bar{Z}_t} w_{\mathrm{ex}}(i,z) w_t^{(i,z)}}$$

with $w_t^{(i,c,z)} = \int \ell_z(x) p_t^{(i,c)}(\mathrm{d}x)$ and $w_t^{(i,z)} = \sum_{c\in C} w_t^{(i,c,z)}$. The only additional parameter is the prior probability for an object to belong to a given class, and this probability distribution over C can simply be defined as uniform.

Pruning can be performed as in the standard case, when the total weight associated with an index in $\hat{\mathbb{I}}_t$ is below the pruning threshold τ_{p}. The merging procedure must however be adapted: distance between hypotheses is obtained by calculating the Hellinger distance [48, Chapt. 3] between the mixtures of the form

$$\sum_{c\in C} \hat{w}_t^{(i,c)} \hat{p}_t^{(i,c)}, \qquad i \in \hat{\mathbb{I}}_t.$$

In consequence, two hypotheses are merged only if all classes have similar weights and probability distributions.

In terms of track extraction, indices in $\hat{\mathbb{I}}_t$ for which the probabilities $\hat{w}_t^{(j,c)}$ sum to more than the given threshold τ_{e} are considered as representing a true object. Then, for each of these selected indices, either there is one $c \in C$ such that $\hat{w}_t^{(j,c)} > \tau'_{\mathrm{e}}$ and the corresponding track is displayed as belonging to the class c, or there is no such class and the track is displayed as a non-classified one.

4.3. Results

The harbour surveillance scenario introduced in Section 3 is considered. A Kalman filter implementation of the HISP filter is considered for addressing the multi-object problem with the

following parameters: the probability of detection is set to 0.75 and the average number of appearing objects is 0.1 per second. The increase of the number of hypotheses is handled by performing pruning with a threshold $\tau_{\mathrm{p}} = 10^{-3}$ and merging with the threshold $\tau_{\mathrm{m}} = 0.5$ for the Hellinger distance. The thresholds considered for track extraction are $\tau_{\mathrm{e}} = 0.9$ and $\tau'_{\mathrm{e}} = 0.99$. In terms of classification, 4 classes are considered based on the dynamics of the objects in the scene:

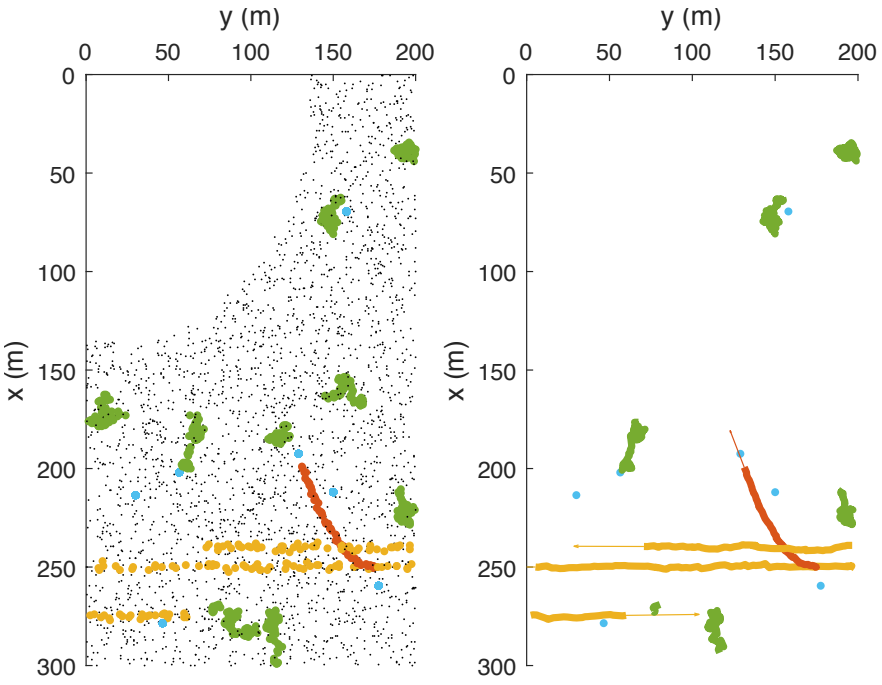
1. Static
2. Moving according to a Brownian motion
3. Moving according to a uniform motion:
   (a) Within the shipping lane
   (b) In the restricted area

Classes 3a and 3b are based on the same dynamics and are distinguished by the value of the x-coordinate of the objects. The static objects of Class 1 are actually modelled by a Brownian motion with small intensity ($0.01\mathrm{m.s}^{-1}$) in order to allow for small displacements. Without this class, static objects in the restricted area would fall in Class 3a which should be kept for suspicious objects only. Class 2 is considered for objects that do not represent a threat, such as schools of fish, and is modelled by a Brownian motion with intensity $0.8\mathrm{m.s}^{-1}$. Finally, the two classes in the third category are associated with a constant-velocity model initialised with an uncertainty of $0.5\mathrm{m.s}^{-1}$ in all directions and a noise of intensity $0.025\mathrm{m.s}^{-2}$.

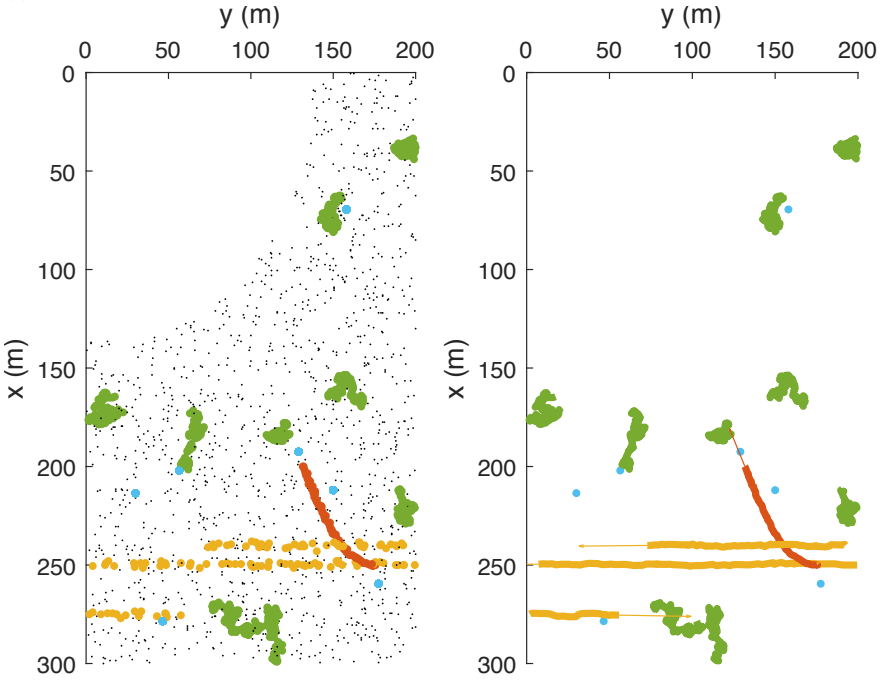
Results are shown in Figure 8 with two different types of seabed: Figure 8a for coarse sand and Figure 8b for muddy sand. In these two cases, the HISP filter tracked the objects of interest and managed to classify them correctly. Because objects are distinguished by the filter, trajectories are also naturally available, so that the AUV can be seen to have been dropped by the second boat in both cases. Velocity vectors are shown for objects in Class 3 only since this component is not estimated in classes based on a Brownian motion. There are visibly more false alarms in the coarse-sand scenario, and the tracking of the schools of fish is made more difficult as a consequence. Indeed, Figure 8a shows that the HISP filter does not manage to recover their full trajectories, which appear truncated when compared with the ground truth. Objects that have a more predictable behaviour, such as static objects and boats, are easier to track and their trajectories is well estimated in both cases. Note that Class 3b is not exactly a "AUV" class, any of the boats would fall in this class if they entered the restricted area. Similarly, the AUV remains in Class 3a as long as it does not cross the boundary of the shipping lane.

In order to assess the performance of the proposed approach, a multi-object performance metric called the OSPA (Optimal Sub-Pattern Assignment) distance [49] has been computed at each time step and for each class as shown in Figure 9. The OSPA distance has been computed with a 2-norm and a cut-off of 50. Results corresponding to a recent version of the PHD filter integrating classification [50] are also indicated for comparison. The HISP filter can be seen to have a better performance than the PHD filter in all classes and in each scenario.

The only exceptions are a) the PHD filter initialises the static targets faster than the HISP filter, which can be due to the fact that the classification of the PHD filter does not have an equivalent to the confirmation threshold τ'_e and therefore confirms a class as soon as the corresponding track is confirmed, and b) the 3 peaks in performance found after 150s for the HISP filter in the boat class, which are due to the fact that the filters maintains track for a few time steps after they leave the surveillance area, hence inducing an error in cardinality. The OSPA distance for the boat class is close to zero for both filters before 70s since no objects with such class are present; the distance



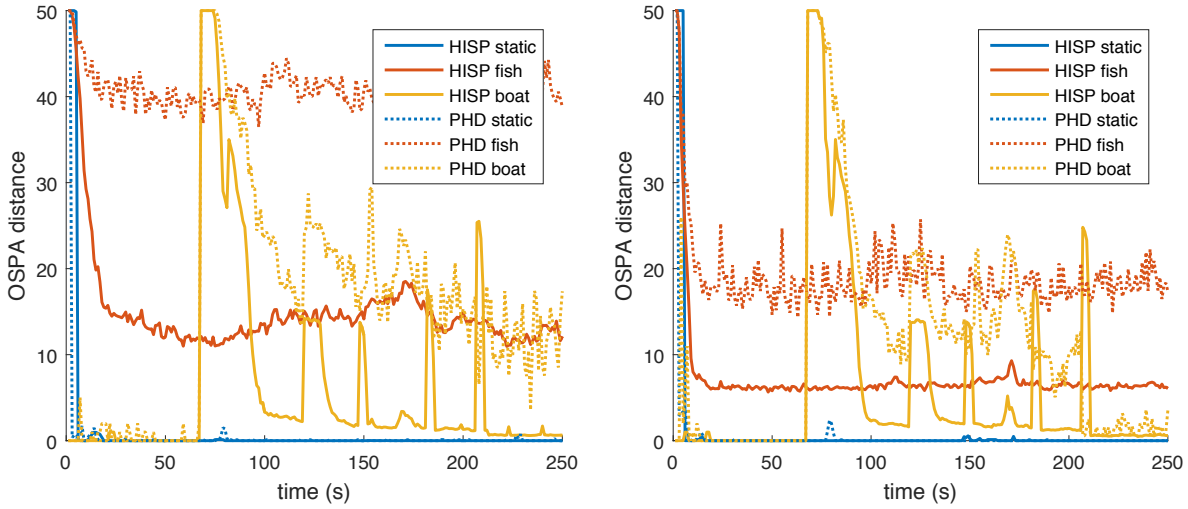
(a) Coarse sand



(b) Muddy sand

Fig. 8: (Colour online) Position, trajectory and velocity of the HISP-filter tracks (right) compared with observations (left, colour-coded according to the ground truth). Colour code: • False alarm – • Fish – • Static object – • Boat – • AUV

is then maximal when the first boat enters the scene as a few time steps are required for the filters to confirm its presence. Overall, the HISP filter appears to be less sensitive to the higher number of false alarms of the coarse-sand scenario since it maintains a similar performance as for the muddy-sand scenario.



(a) Coarse sand (b) Muddy sand

Fig. 9: OSPA distance per class (averaged over 50 Monte Carlo runs)

5. ACOUSTICAL TRACKER

In Section 4, a digital tracker based on detections of the fused MIMO data was developed. This section shows how tracking can be performed automatically at a sensor level using time reversal techniques. Thanks to its multiple-sensor structure, MIMO sonar systems are particularly well suited for time reversal techniques. To implement Time Reversal techniques, each element of the MIMO sonar system is both a transmitter and a receiver.

First developed by Fink [30], such techniques utilise the fact that the solution of the wave equation is invariant under time reversal, *i.e.* if $p(t)$ is solution of the wave equation then $p(-t)$ is also solution. Prada et al. in [31] described the iterative time reversal process for a static scene. The MIMO problem formulation can be written as:

$$\mathbf{R}(\omega) = \mathbf{K}(\omega)\mathbf{E}(\omega) \tag{19}$$

where $\mathbf{E}(\omega)$ is the column vector of the Fourier transform of the transmit signals, $\mathbf{R}(\omega)$ is the column vector of the Fourier transform of the received signals and $\mathbf{K}(\omega)$ the channel matrix. So in the time domain, the received signals are the convolutions between the pulses and the channel matrix. In her paper, Prada considered an interactive version of time reversal, meaning that the received signals are time reversed before being sent back into the environment. At iteration n and given a received signal $\mathbf{R}_n(\omega)$, the next output signals are then given by:

$$\mathbf{E}_{n+1}(\omega) = \mathbf{R}_n^*(\omega) = \mathbf{K}^*(\omega)\mathbf{E}_n^*(\omega) \tag{20}$$

Using this formulation and collocated transmitters and receivers, the $2n^{\text{th}}$ input signals is:

$$\mathbf{E}_{2n}(\omega) = [\mathbf{K}^*(\omega)\mathbf{K}(\omega)]^n \mathbf{E}_0(\omega) \tag{21}$$

Prada shows that the $[\mathbf{K}^*(\omega)\mathbf{K}(\omega)]^n$ operator converges to the brightest scattering point of the scene. Effectively the MIMO array focuses the sound to this scattering point.

Considering now a dynamic scene, such as the one described in Section 3, the channel matrix varies with time and can be written as $\mathbf{K} = \mathbf{K}(\omega, t)$. The environment is sensed at discrete time step $\{t_n\}_{n\in\mathbb{N}}$ by the MIMO sonar system. For the sake of the discussion, the channel matrix is assumed constant during the MIMO sensing process. Let $\mathbf{K}_n(\omega) = \mathbf{K}(\omega, t_n)$ be the channel matrix at time step t_n. Equation (19) leads to:

$$\mathbf{R}_n(\omega) = \mathbf{K}_n(\omega)\mathbf{E}_n(\omega) \tag{22}$$

Note that $\mathbf{K}_n(\omega)$ can be divided into two terms: a static component $\mathbf{K}^s(\omega)$ coming from the static components of the scene (e.g. seabed) and a dynamic component $\mathbf{K}_n^d(\omega)$ representing the moving objects in the scene. $\mathbf{K}^s(\omega)$ can be learned and suppressed from equation (22) thanks to background subtraction: always transmitting the same pulse $\mathbf{E}_0$ for a period of time leads to:

$$\mathbf{R}_n(\omega) = \mathbf{K}^s(\omega)\mathbf{E}_0(\omega) + \mathbf{K}_n^d(\omega)\mathbf{E}_0(\omega) \tag{23}$$

Assuming that $\mathbb{E}[K_n^d(\omega)] = 0$ (where $\mathbb{E}[.]$ represents the expectation), Eq. (23) leads to:

$$\mathbb{E}[\mathbf{R}_n(\omega)] = \mathbf{K}^s(\omega)\mathbf{E}_0(\omega) \tag{24}$$

Equation (22) can then be rewritten as:

$$\mathbf{R}_n(\omega) - \mathbb{E}[\mathbf{R}_n(\omega)] = \mathbf{K}_n^d(\omega)\mathbf{E}_n(\omega) \tag{25}$$

For the sake of the discussion and readability of the equations, the static component has been removed in the followings equations.

In order to track an underwater target in motion, the proposed approach defocuses the input signal $\mathbf{E}_{n+1}$ using the defocus function $\mathbf{G}$ accordingly to the maximum speed of the target and the pulse repetition rate. $\mathbf{E}_{n+1}$ then becomes $\mathbf{E}_{n+1} = \mathbf{G}\mathbf{K}_n^*\mathbf{E}_n^*$, and equation (21) can be rewritten as:

$$\mathbf{E}_{2N} = \left[\prod_{2n=2}^{2N} \mathbf{G}\mathbf{K}_{2n-1}^*\mathbf{G}^*\mathbf{K}_{2n-2}\right]\mathbf{E}_0 \tag{26}$$

Note that the ω variable has been omitted for readability. Figure 10 displays the procedure for acoustical tracking using a MIMO sonar system linking each step to its relative equation. After the transmission of the initial pulse $\mathbf{E}_0$, the MIMO echoes $\mathbf{R}_0$ corresponding to Eq. (22) are recorded. The static component of the echoes are removed according to Eq. (25). Finally the time reversed echoes are sent back into the environment as stipulates Eq. (26).

Figure 11 shows a comparison between classical Time Reversal processing and the proposed iterative unfocused Time Reversal technique. A moving target T is located at the position $[x = 100\text{m}, y = 200\text{m}]$ at the time $t = t_n$. At $t = t_{n+1}$, T is located at $[x = 98.5\text{m}, y = 200\text{m}]$. Figure 11 displays the normalised energy of the back-propagated total acoustical field received at time $t = t_n$. As expected, using the classical implementation of the Time Reversal technique proposed by [31] focuses the acoustical energy on the target present at t_n. As pictured in figure 11(a), the focused

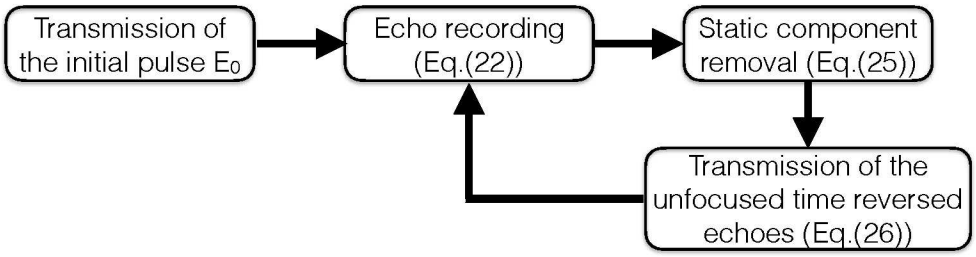
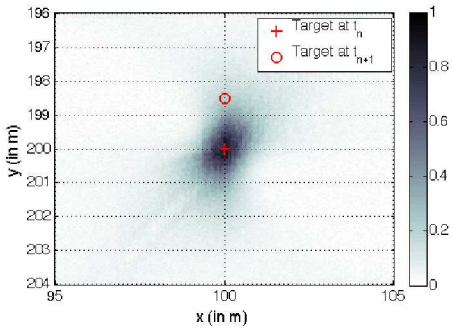
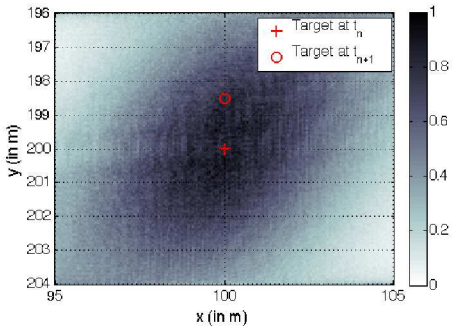


Fig. 10: Unfocused Time Reversal procedure.



(a)



(b)

Fig. 11: Normalised energy of the back-propagated acoustical field using (a) traditional Time Reversal (Eq. (21)), and (b) unfocused Time Reversal (Eq. (26)).

energy missed the target at t_{n+1}. By introducing the defocus function in Eq. (26), the target T at t_n is still in the blob of energy back-propagated from the received echo $\mathbf{R}_n$.

It is interesting to note that the iterative defocussed time reversal process follows in essence the general approach taken by digital tracking filters. Tracking algorithms indeed proceed in two steps:

$$p_k(X_k|Z^{(k)}) \rightarrow p_{k+1|k}(X_{k+1}|Z^{(k)}) \rightarrow p_{k+1}(X_{k+1}|Z^{(k+1)}) \tag{27}$$

The first step is a prediction step and is equivalent to the defocus operator $\mathbf{G}$. As its name suggests, the prediction step infers the position of the target for the next time step and also propagates the uncertainty of this prediction. Defocusing the signal accordingly to the target maximum speed

ensures that the most acoustical energy will be sent toward the target at the next time step. The second step is the data update is equivalent to the channel matrix operator $\mathbf{K}_n$. During the data update, the state space is updated using the actual observations. In the acoustical tracker, the targets' positions are automatically updated by sensing the environment at the next time step.

6. Conclusion

This paper proposed a practical application of distributed sensors to the difficult problem of harbour protection. It had been shown that static multi-distributed MIMO sonar systems offer an interesting, versatile and reliable platform for area surveillance. In particular, this paper has demonstrated great capabilities of such system to track and classify low profile mid-water target such as AUV or diver. A realistic full 3D MIMO sonar simulator was developed. The simulator can compute efficiently the full MIMO sonar system response for complex environments and offer realistic synthetic data to work on. The efficiency of a novel filter scheme, the HISP, has been discussed and applied successfully to MIMO data. It had been shown that the HISP filter enabled the tracking and classification of the objects in the scene and an automatic detection of potential threats. Alternatively, the multiplicity of sensors of a large MIMO sonar systems brings degrees of freedom to the way one can exploit such system. In this context, the founding principles of a fully automatic acoustical tracker based on time-reversal principles and on the defocusing of the output signals has been proposed. The proposed digital and acoustical trackers address different aspects of the problem and it would be possible, and indeed natural, to attempt to combine them by generalising the observation process of the HISP filter to include time-dependent observations. Although out of scope for the current article, this is a promising research direction that will be considered for future works.

Acknowledgment

This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) Grant number EP/J015180/1 and the MOD University Defence Research Collaboration in Signal Processing EP/K014277/1.

7. References

[1] Kessel, R. T., Hollett, R. D.: 'Underwater intruder detection sonar for harbour protection: State of the art review and implications', Tech. Rep. NURC-PR-2006-027, (NATO Undersea Research Centre, 2006)

[2] Kessel, R. T.: 'NATO harbour protection trials 2006 (HPT06) analyst report: Harbour surveillance systems', Tech. Rep. NURC-FR-2007-004, (NATO Undersea Research Centre, 2007)

[3] Juan, Y., Feng, X., Jia, L., Xudong, A., Yongqiang, J.: 'The reverberation suppression in wideband diver detection sonar', Proc. IEEE OCEANS Conf., St. Johns, Sept 2014, pp. 1–4

[4] Xinke, L., Zhengxiang, X.: 'Underwater small target tracking algorithm based on diver detection sonar image sequences', Proc. Industrial Control and Electronics Engineering (ICICEE), Aug 2012, pp. 727–730

[5] DeMarco, K. J., West, M. E., Howard, A. M.: 'Sonar-based detection and tracking of a diver for underwater human-robot interaction scenarios', Proc. 2013 IEEE International Conference on Systems, Man, and Cybernetics, Oct 2013, pp. 2378–2383

[6] Chen, X., Wang, R., Tureli, U.: 'Passive acoustic detection of divers under strong interference', Proc. IEEE OCEANS Conf., Sept 2006, pp. 1–6

[7] Stolkin, R., Florescu, I.: 'Probabilistic analysis of a passive acoustic diver detection system for optimal sensor placement and extensions to localization and tracking', Proc. IEEE OCEANS Conf., Sept 2007, pp. 1–6

[8] Fillinger, L., Hunter, A. J., Zampolli, M., Clarijs, M. C.: 'Passive acoustic detection of closed-circuit underwater breathing apparatus in an operational port environment', J. Acoust. Soc. Amer., 2012, 132, (4), EL310–EL316

[9] Sharma, N. S., Yakubovskiy, A. M., Zimmerman, M. J.: 'Scuba diver detection and classification in active and passive sonars; a unified approach', Proc. Technologies for Homeland Security (HST), Nov 2013, pp. 189–194

[10] Hari, V. N., Chitre, M., Too, Y. M., Pallayil, V.: 'Robust passive diver detection in shallow ocean', Proc. IEEE OCEANS Conf., Genova, May 2015, pp. 1–6

[11] Kim, S., Ku, B., Hong, W., Ko, H.: Performance comparison of target localization for active sonar systems', Aerospace and Electronic Systems, IEEE Transactions on, 44, (4), 2008, pp. 1371–1380

[12] Fewell, M., Ozols, S.: 'Simple detection-performance analysis of multistatic sonar for anti-submarine warfare'. Tech. rep., (DSTO Defence Science and Technology Organisation, 2011)

[13] Ehlers, F.: 'Final report on deployable multistatic sonar systems', Tech. rep., (NATO Undersea Research Centre, 2009)

[14] Orlando, D., Ehlers, F. 'Advances in multistatic sonar', in Kolev, N. Z. (Ed.): 'Sonar Systems', (InTech, 2011), pp. 29–50

[15] Pailhas, Y., Petillot, Y., Mulgrew, B., Brown, K.: 'Spatially distributed MIMO sonar systems: principles and capabilities' Oceanic Engineering, IEEE Journal of, *accepted*, 2016

[16] Pailhas, Y., Houssineau, J., Delande, E., Petillot, Y., Clark, D. E.: 'Tracking underwater objects using large MIMO sonar systems', Proc. Int. Conf. on Underwater Acoustics (UA2014), Rhodes, Greece, June 2014, pp. 1033–1040

[17] Pailhas, Y. Petillot, Y.: 'Large MIMO sonar systems: A tool for underwater surveillance', Proc. Sensor Signal Processing for Defence (SSPD), Edinburgh, UK, Sept 2014, pp. 1–5

[18] Blackman, S. S.: 'Multiple hypothesis tracking for multiple target tracking', Aerospace and Electronic Systems Magazine, IEEE Transactions on, 2004, 19, (1), pp. 5–18

[19] Mahler, R. P. S.: 'Multitarget bayes filtering via first-order multitarget moments', Aerospace and Electronic Systems, IEEE Transactions on, 2003, 39, (4), pp. 1152–1178

[20] Mahler, R. P. S.: 'PHD filters of higher order in target number', Aerospace and Electronic Systems, IEEE Transactions on, 2007, 43, (4), pp. 1523–1543

[21] Erkmen, B. Yildirim, T.: 'Improving classification performance of sonar targets by applying general regression neural network with PCA', Expert Systems with Applications, 2008, 35, (2), pp. 472–475

[22] Shams, S.: 'Neural network optimization for multi-target multi-sensor passive tracking', Proceedings of the IEEE, 1996, 84, (10), pp. 1442–1457

[23] Ayrulu, B., Barshan, B. 'Neural networks for improved target differentiation and localization with sonar', Neural Networks, 2001, 14, (3), pp. 355 – 373, 2001

[24] Perry, S. W., Guan, L.: 'A recurrent neural network for detecting objects in sequences of sector-scan sonar images', IEEE Journal of Oceanic Engineering, 2004, 29, (3), pp. 857–871

[25] Smith, D., Singh, S.: 'Approaches to multisensor data fusion in target tracking: A survey', IEEE Transactions on Knowledge and Data Engineering, 2006, 18, (12), pp. 1696–1710

[26] Winter, M., Favier, G.: 'A neural network for data association', Proc. Acoustics, Speech, and Signal Processing, Phoenix, U.S.A, Mar 1999, pp. 1041–1044

[27] Houssineau, J.: 'Representation and estimation of stochastic populations', PhD thesis, Heriot-Watt University, 2015

[28] Houssineau, J., Clark, D. E., Del Moral, P.: 'A sequential Monte Carlo approximation of the HISP filter', Proc. European Signal Processing Conference (EUSIPCO), Nice, France, 2015

[29] McKenna, I., Tonolini, F., Tobin, R., *et al.*: 'Observing the dynamics of waterborne pathogens for assessing the level of contamination', Proc. Sensor Signal Processing for Defence Conf., Edinburgh, UK, 2015

[30] Fink, M., Prada, C., Wu, F., Cassereau, D.: 'Self focusing in inhomogeneous media with time reversal acoustic mirrors', Proc. Ultrasonics Symposium, Oct 1989, pp. 681–686

[31] Prada, C., Thomas, J., Fink, M.: 'The iterative time reversal process: Analysis of the convergence', J. Acoust. Soc. Am., 1994, 97, (1), pp. 62–71

[32] Song, H., Kuperman, W., Hodgkiss, W., Akal, T., Guerrini, P.: 'Demonstration of a high-frequency acoustic barrier with a time-reversal mirror', IEEE Journal Oceanic Engineering, 2003, 28, (2), pp. 246–249

[33] Sabra, K. G., Roux, P., Song, H.-C., Hodgkiss, W. S., Kuperman, W. A., Akal, T., Stevenson, J. M.: 'Experimental demonstration of iterative time-reversed reverberation focusing in a rough waveguide. application to target detection', J. Acoust. Soc. Am., 2006, 120, (3), pp. 1305–1314

[34] Roux, P., Akal, T., Hodgkiss, W. S., Kuperman, W. A., Song, H. C., Stevenson, J. M.: Time-reversal arrays in underwater acoustics. J. Acoust. Soc. Am., 2004, 116, (4), pp. 2526–2526

[35] Mandelbrot, B.: 'The fractal geometry of nature', (W H Freeman & Co, 1982)

[36] Kaplan, L., Kuo, C.-C.: 'An improved method for 2-D self-similar image synthesis', Image Processing, IEEE Transactions on, 1996, pp. 754–761

[37] Williams, K., Jackson, D.: 'Bistatic bottom scattering: Model, experiments, and model/data comparison', Tech. rep., (APL-UW, 1997)

[38] Etter, P. C.: 'Underwater Acoustic Modeling and Simulation', (Spon Press, 2003)

[39] de Sousa Costa, E., Bauzer Medeiros, E., Carvalho Filardi, J. B.: 'Underwater acoustics modeling in finite depth shallow waters', in Beghi, M. G. (Ed): 'Modeling and Measurement Methods for Acoustic Waves and for Acoustic Microdevices', (InTech, 2013)

[40] Porter, M. B., Bucker, H. P.: 'Gaussian beam tracing for computing ocean acoustic fields', J. Acoust. Soc. Am., 1987, 82, (4), pp. 1349–1359

[41] Lurton, X.: 'An Introduction to Underwater Acoustics, Principles and Applications', (Springer-Verlag Berlin Heidelberg, 2010)

[42] Pailhas, Y., Petillot, Y., Capus, C., Mulgrew, B.: 'Target detection using statistical MIMO', Proc. of Meetings on Acoustics, Edinburgh, UK, 2012

[43] Home, A., Yates, G.: 'Bistatic synthetic aperture radar', IEEE RADAR, 2002, pp. 6–10

[44] Panta, K., Vo, B.-N., Singh, S., and Doucet, A.: 'Probability hypothesis density filter versus multiple hypothesis tracking', Defense and Security, International Society for Optics and Photonics, 2014, pp. 284–295

[45] Clark, D. E., Panta, K., Vo, B.-N.: 'The GM-PHD filter multiple target tracker', Proc. Information Fusion, 2006 9th International Conference on, 2006, pp. 1–8

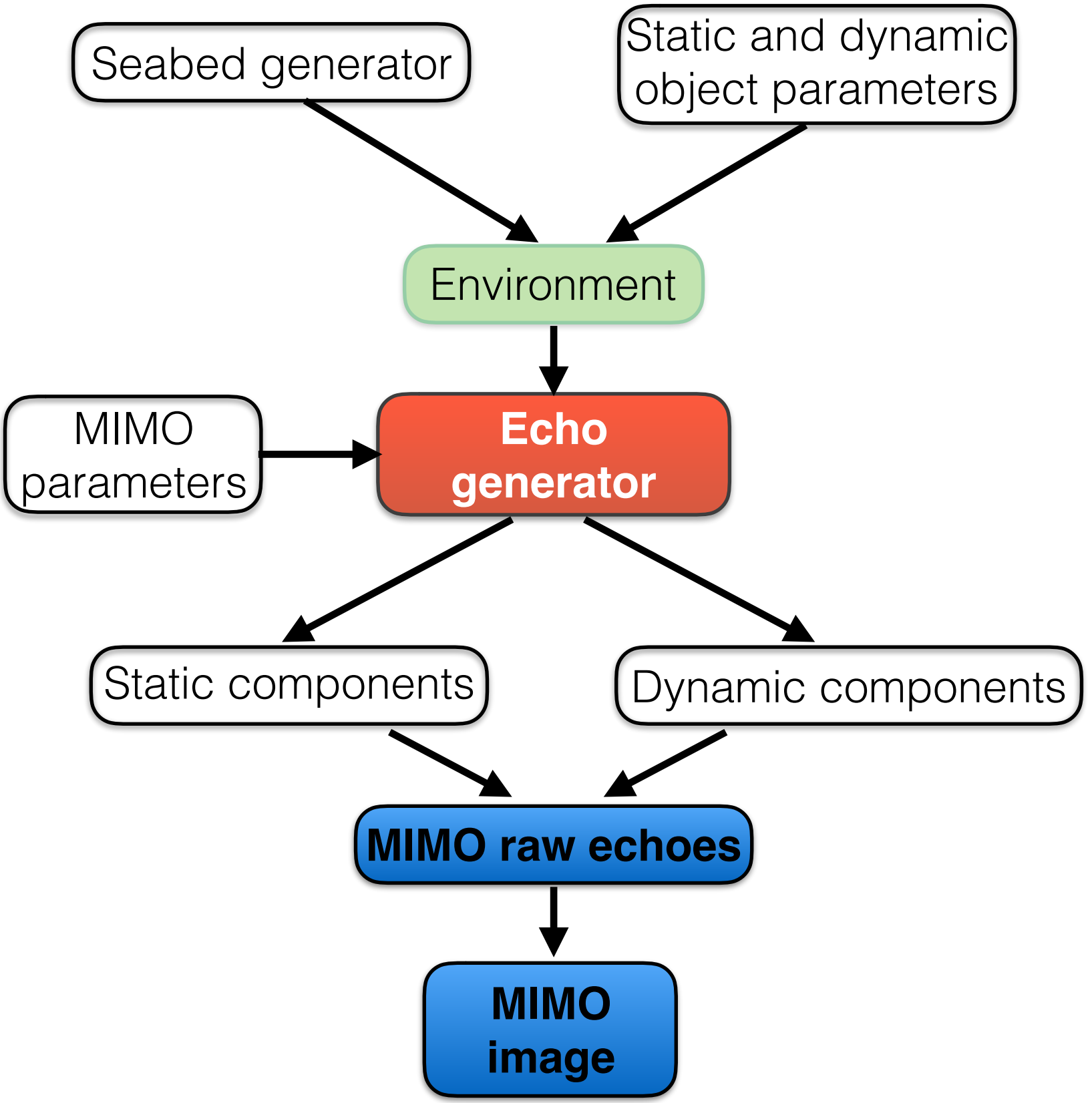
[46] Clark, D. E., Bell, J.: 'Bayesian multiple target tracking in forward scan sonar images using the phd filter', IET Radar, Sonar and Navigation, 2005, 152, pp. 327–334

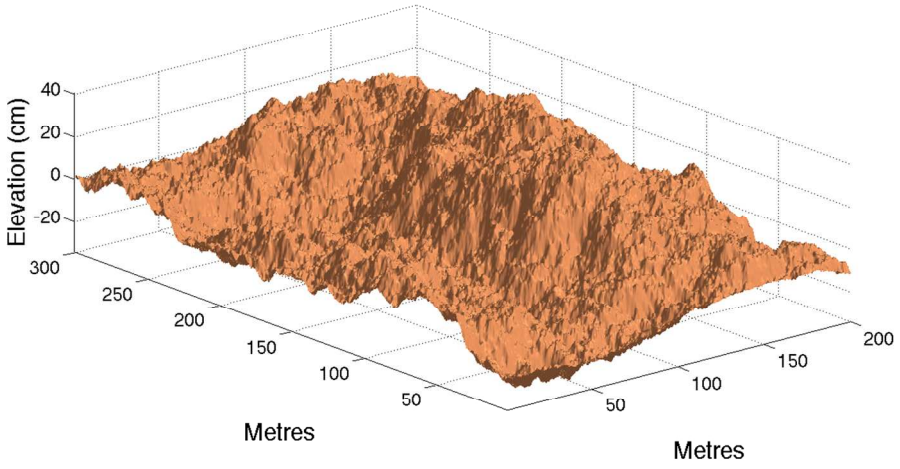
[47] Vo, B.-N., Ma, W.-K.: 'The Gaussian mixture probability hypothesis density filter', Signal Processing, IEEE Transactions on, 2006, 54, (11), pp. 4091–4104

[48] Pollard, D.: 'A user's guide to measure theoretic probability', (Cambridge University Press, vol. 8., 2002)

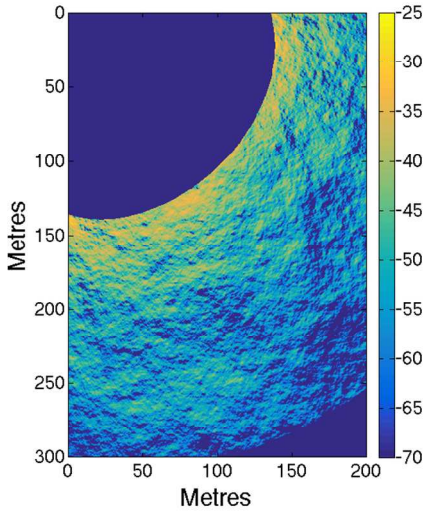
[49] Schuhmacher, D., Vo, B.-T., Vo, B.-N.: 'A consistent metric for performance evaluation of multi-object filters', Signal Processing, IEEE Transactions on, 2008, 56, (8), pp. 3447–3457

[50] Schlangen, I., Franco, J., Houssineau, *et al.*: 'Marker-less stage drift correction in super-resolution microscopy using the single-cluster PHD filter', Selected Topics in Signal Processing, IEEE Journal of, 2016, 10, (1), pp. 193–202

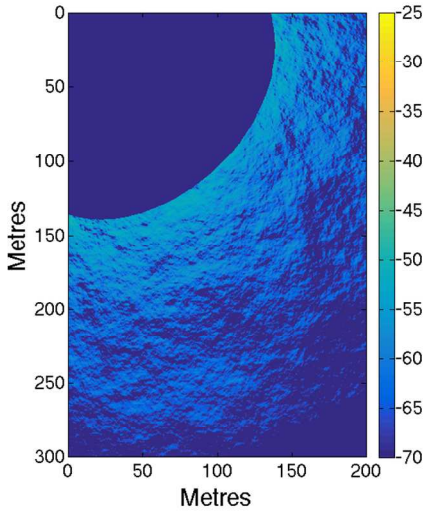

Seabed generator
Static and dynamic object parameters
Environment
MIMO parameters
Echo generator
Static components
Dynamic components
MIMO raw echoes
MIMO image



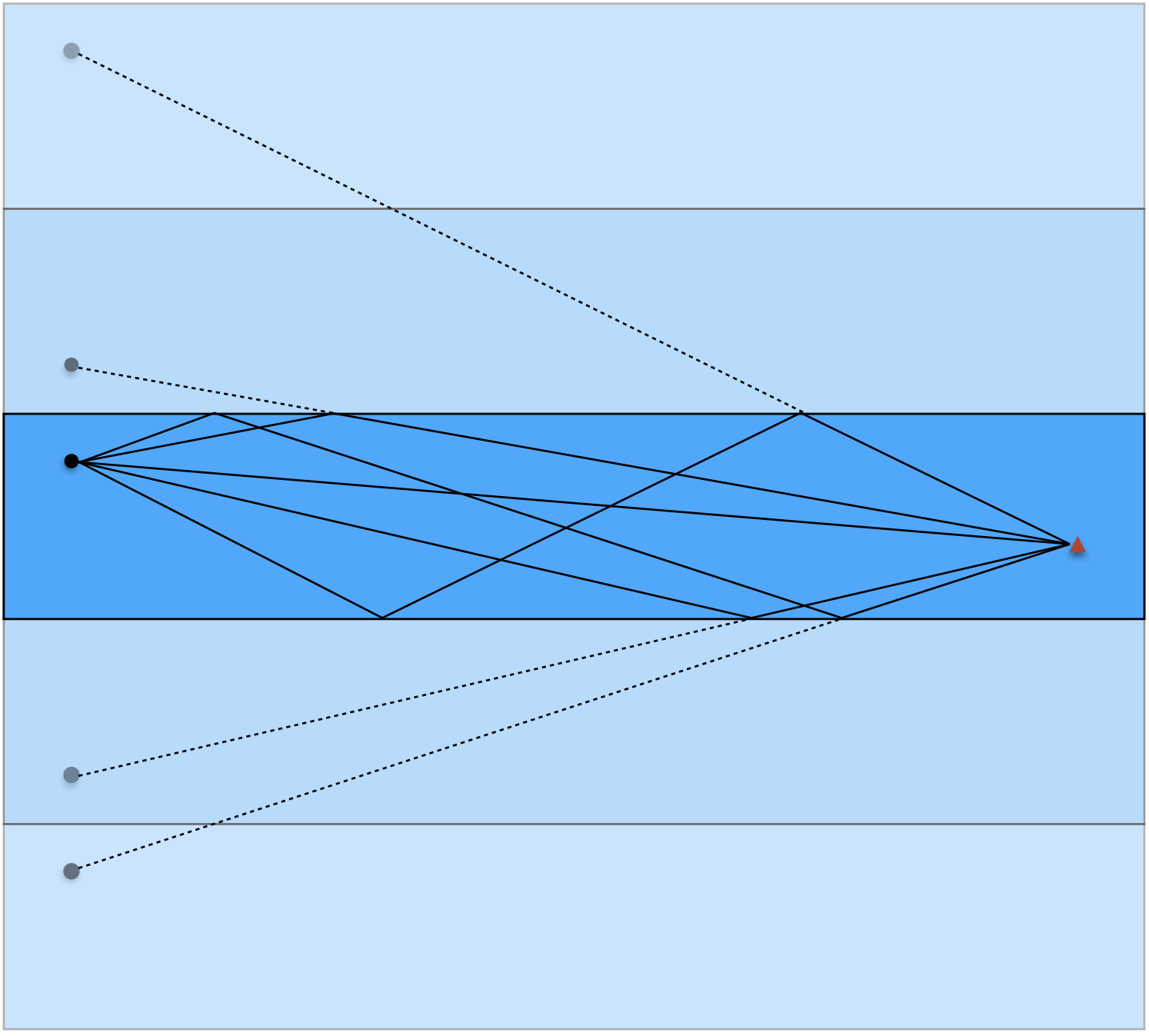
(Colour online) Example of 2D fBm with H = 0.8 (fractal dimension = 2.2)

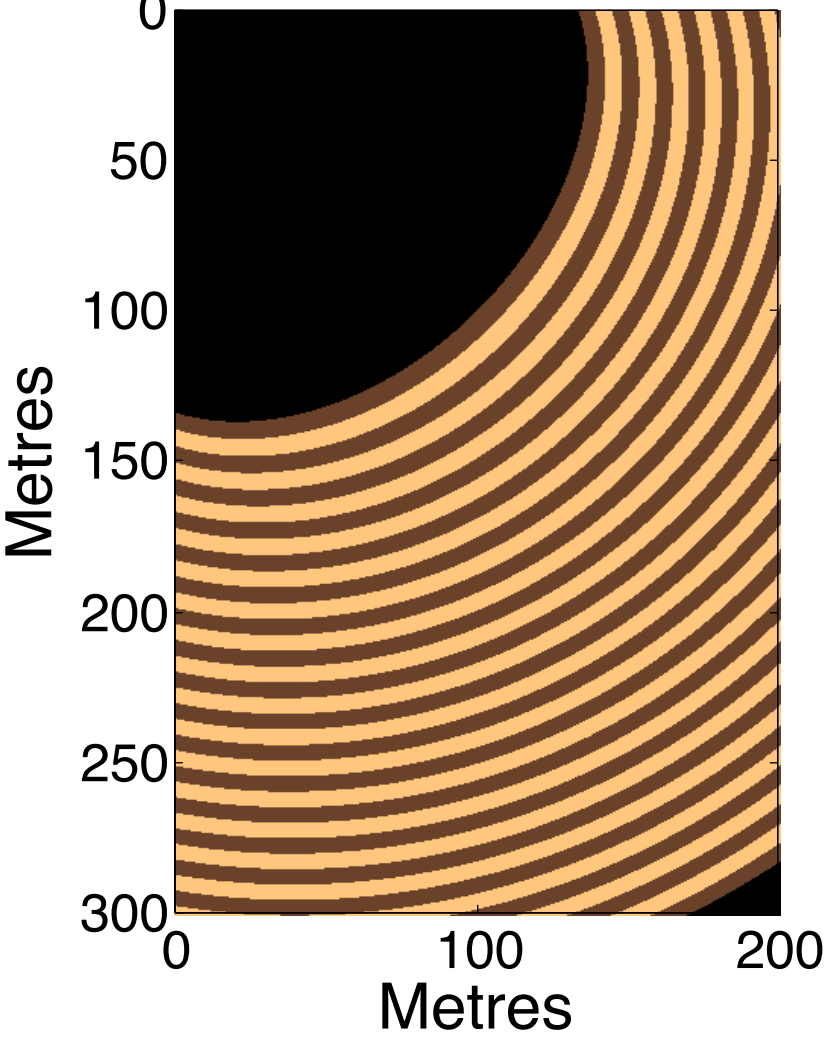


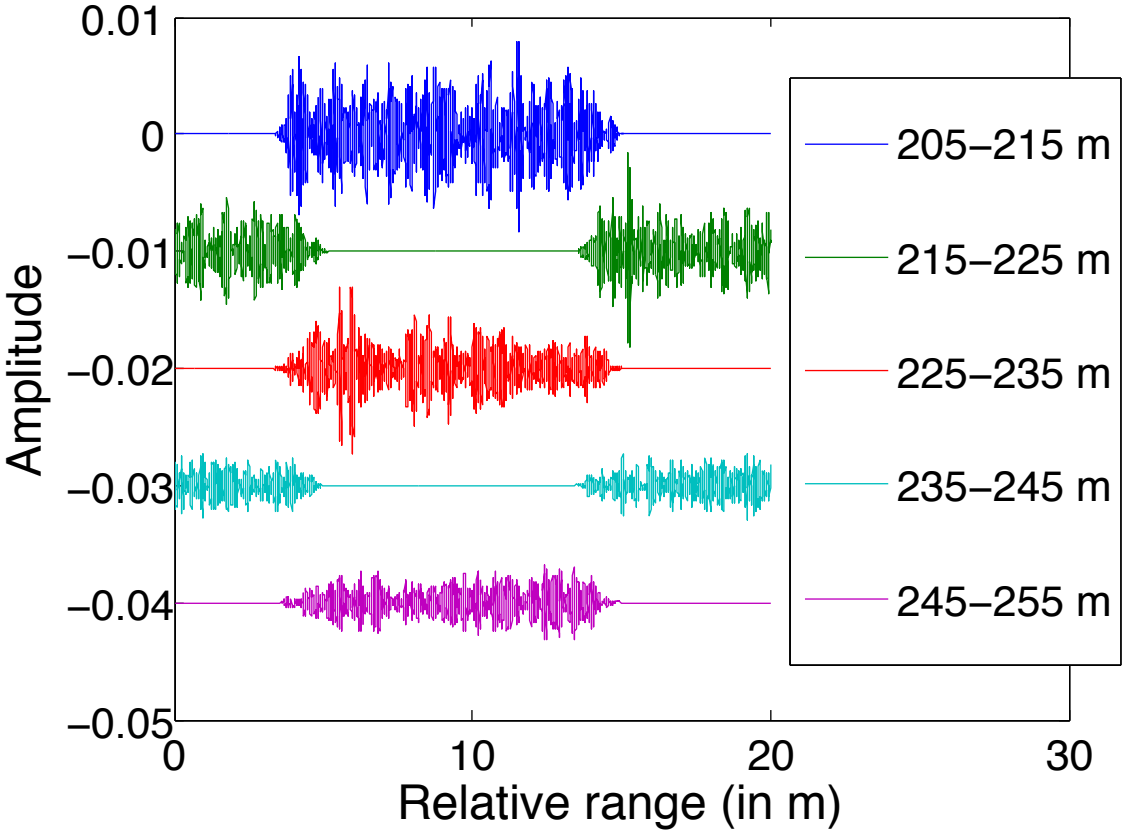
(Colour online) Bistatic scattering strength in dB relative to one Tx located at [x=0m, y=100m, z=7.5m] and a Rx located at [x=100m, y=0m, z=7.5m] for (a) a coarse sand sediment type and (b) a sandy mud sediment type.

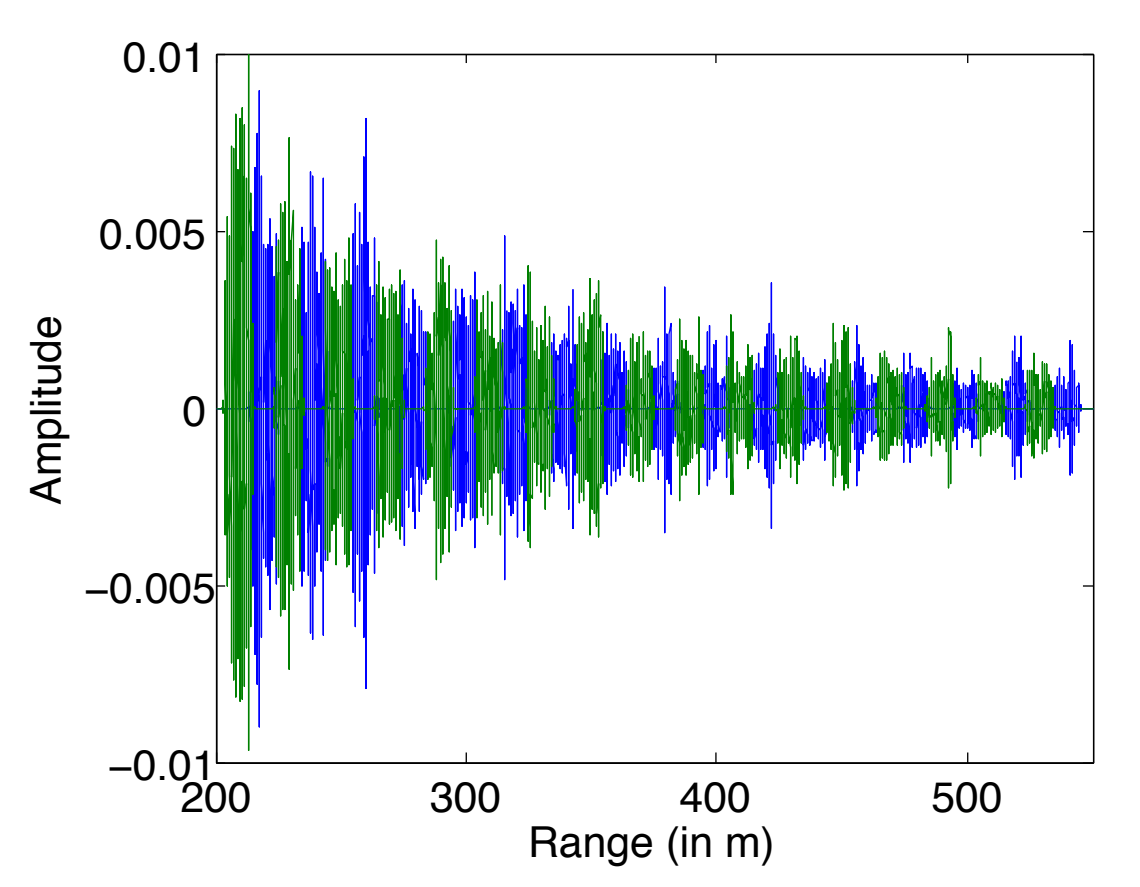


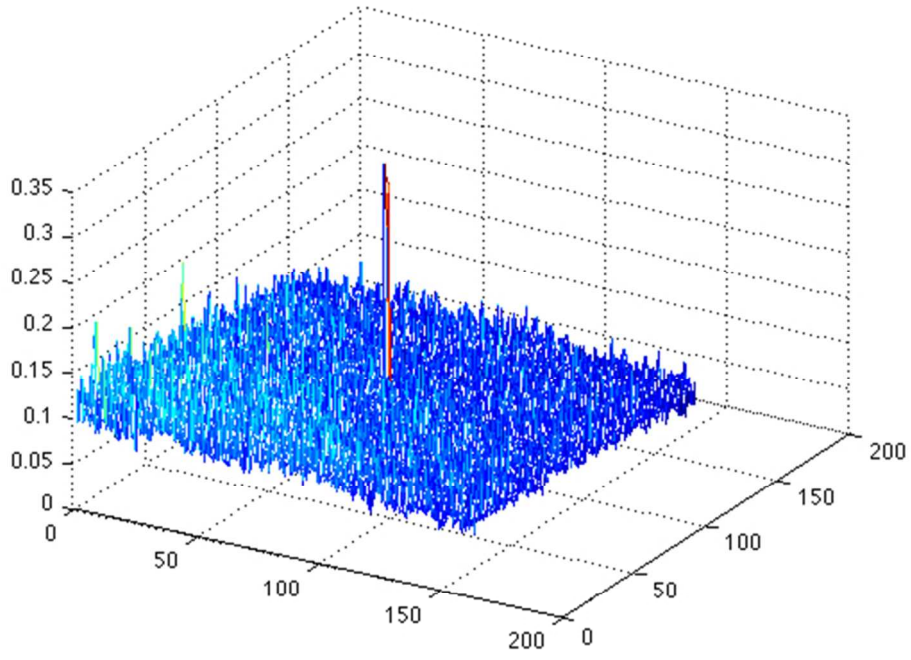
(Colour online) Bistatic scattering strength in dB relative to one Tx located at [x=0m, y=100m, z=7.5m] and a Rx located at [x=100m, y=0m, z=7.5m] for (a) a coarse sand sediment type and (b) a sandy mud sediment type.




0
50
100
150
200
250
300
Metres
0
100
200
Metres

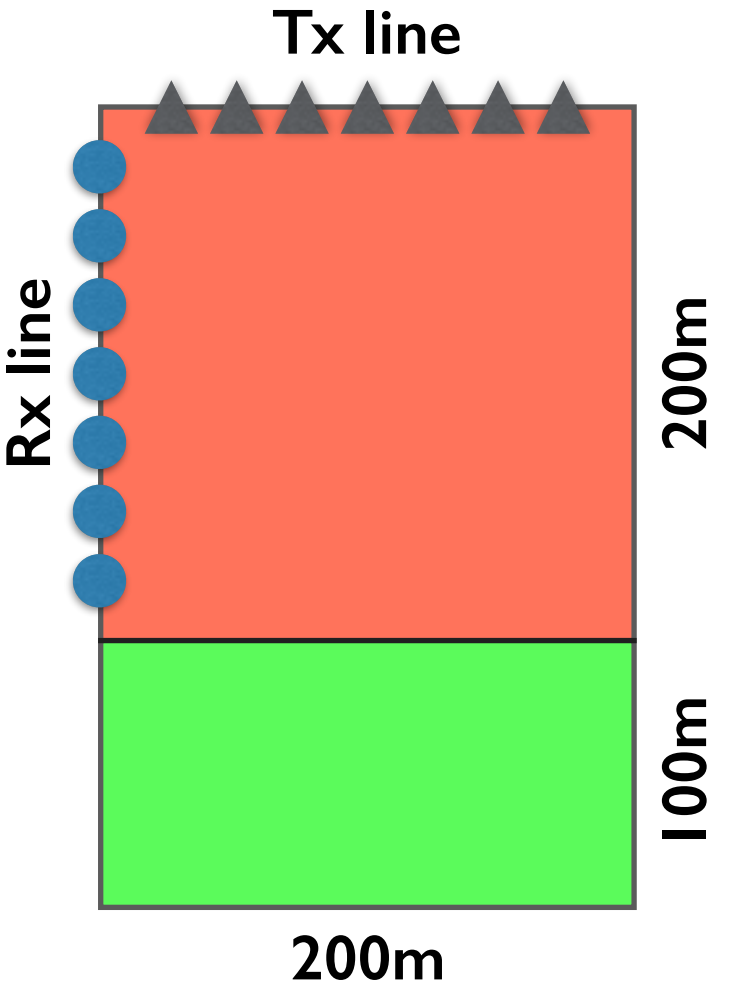

0.01
0
−0.01
−0.02
−0.03
−0.04
−0.05
Amplitude
0
10
20
30
Relative range (in m)
205–215 m
215–225 m
225–235 m
235–245 m
245–255 m

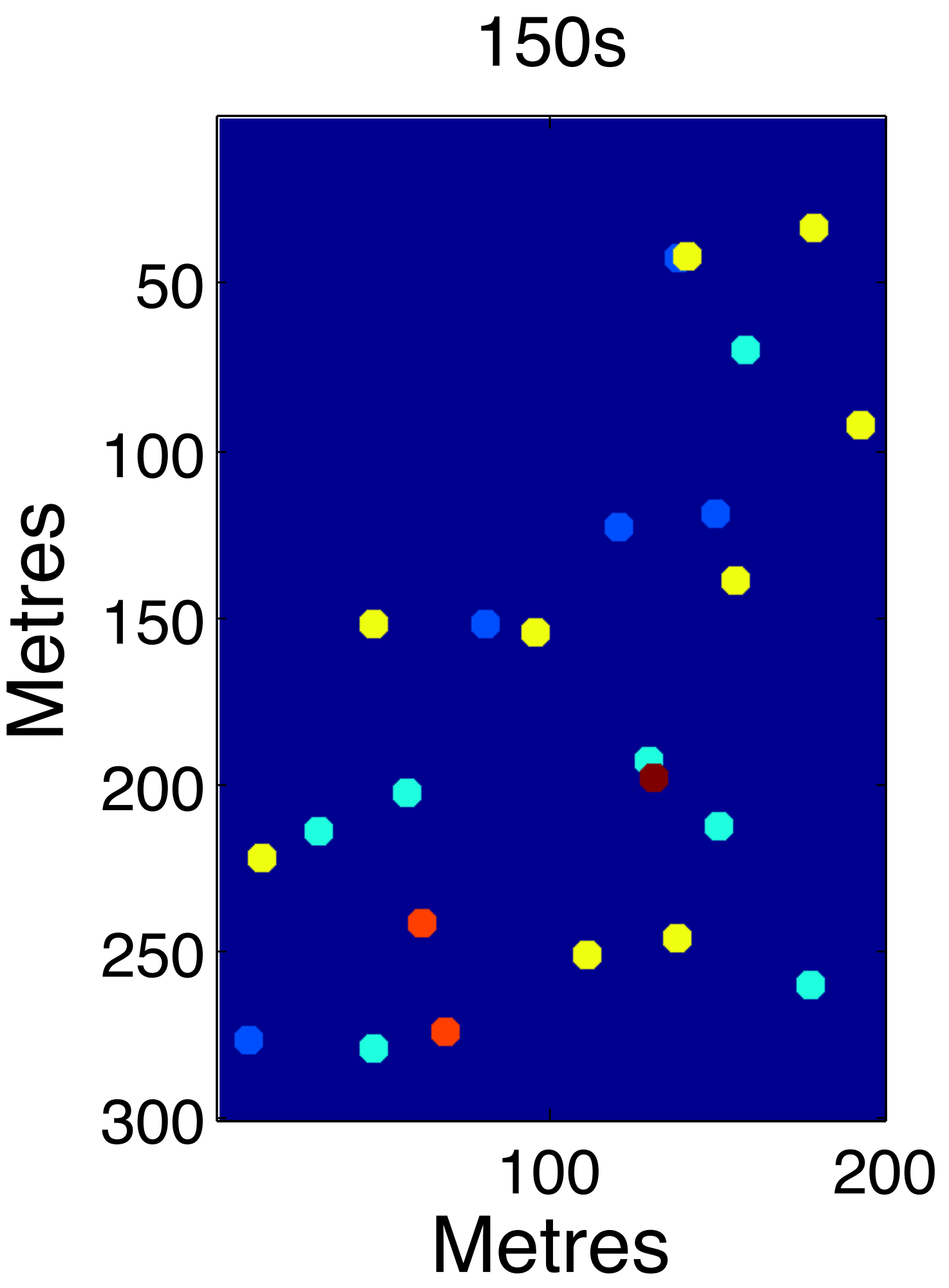

0.01
0.005
0
−0.005
−0.01
Amplitude
200
300
400
500
Range (in m)

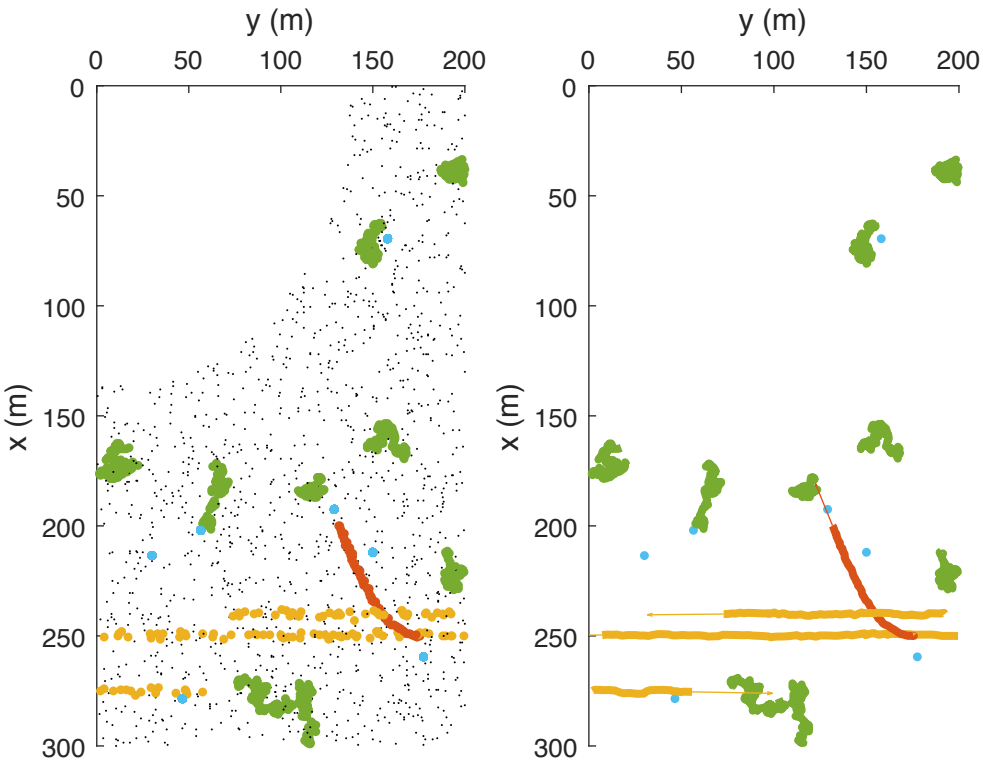


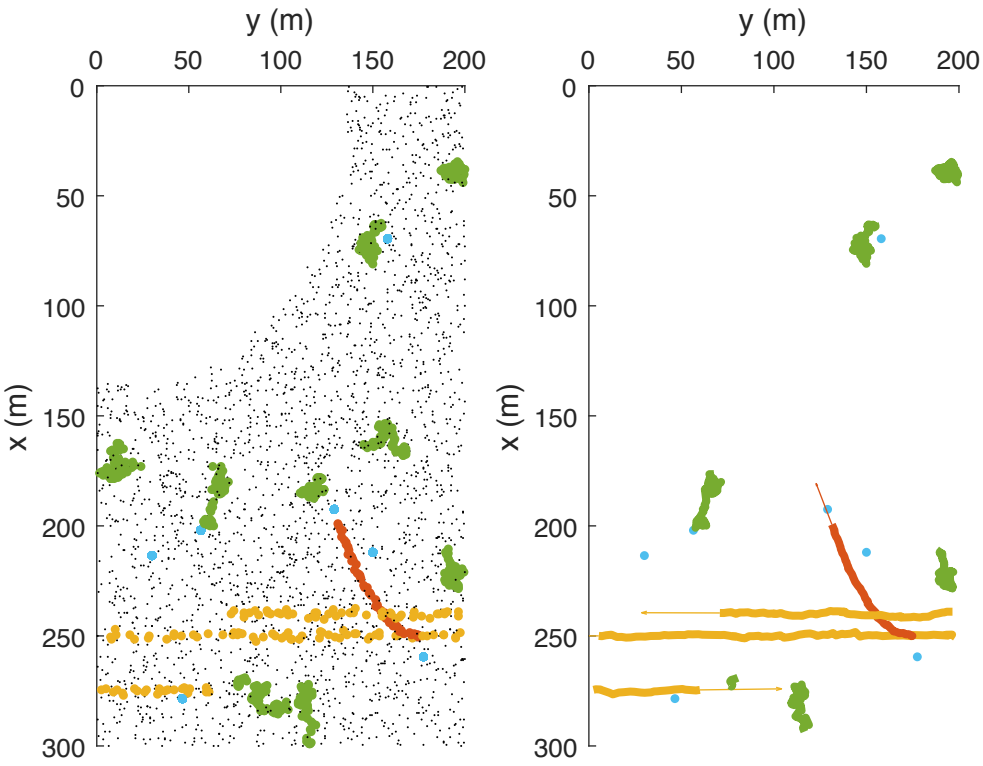
(Colour online) Surface plot of the synthetic aperture MIMO image of a mid water -30dB target on a coarse sand sediment background~\cite{Pailhas2014}.

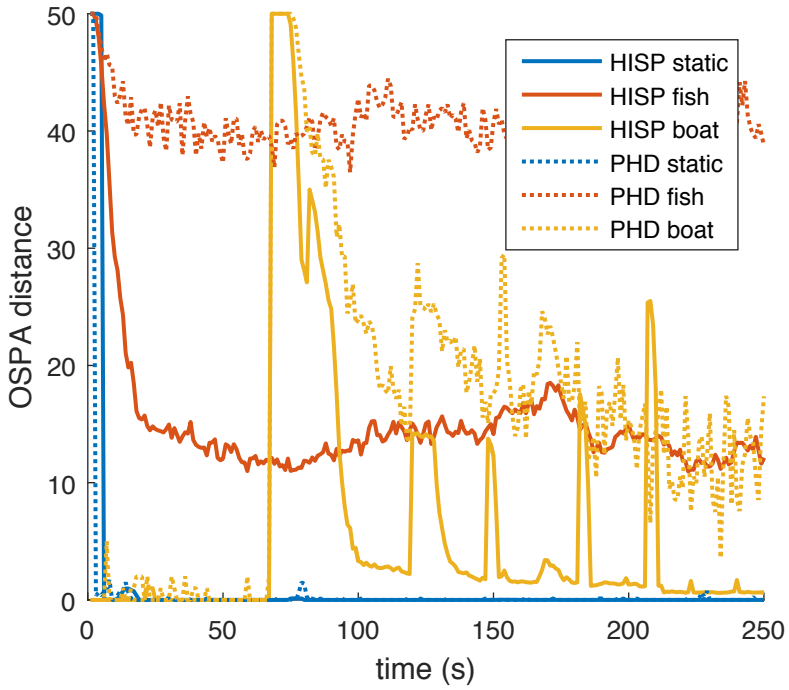

Restricted area
Heavy traffic
Torontohaven
Botlek
Welplaathaven
BOTLEK

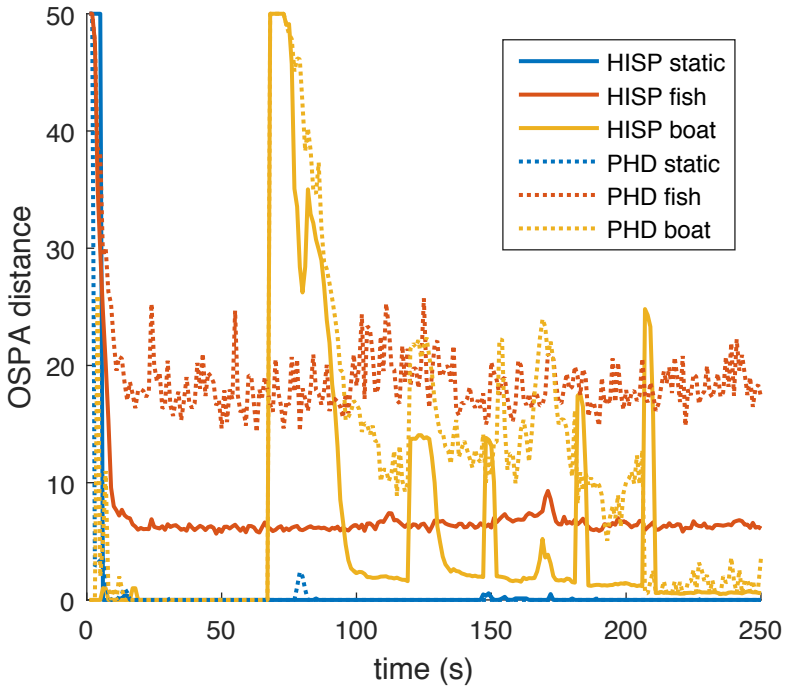

Tx line
Rx line
200m
100m
200m

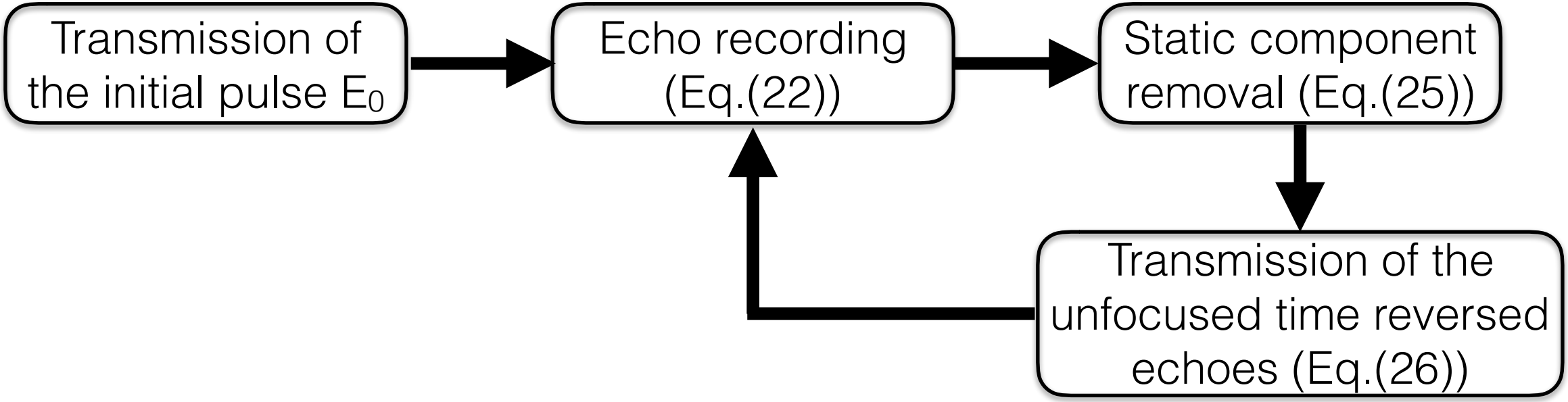

150s
50
100
150
200
250
300
Metres
100
200
Metres

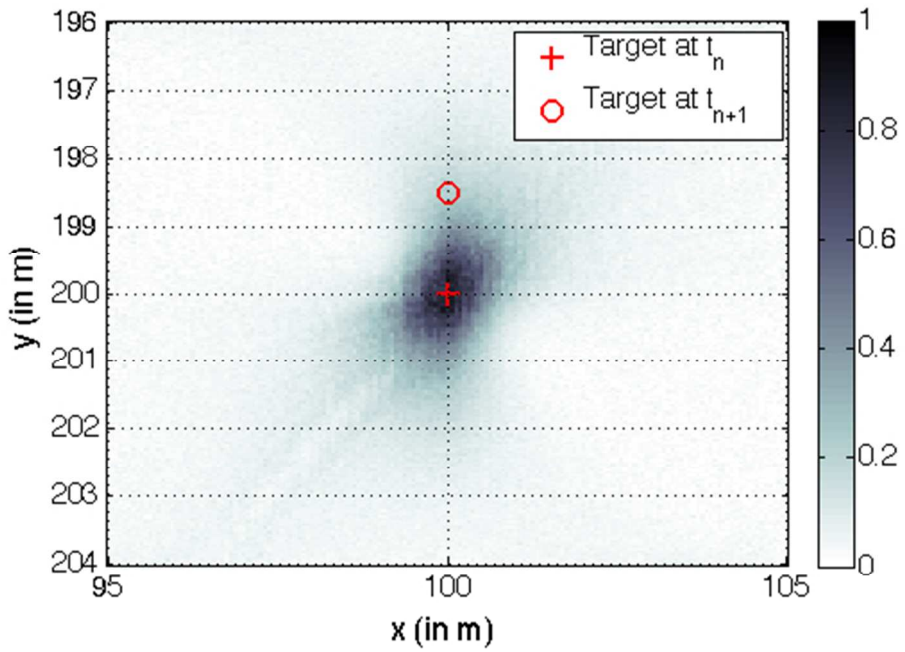

y (m)
0
50
100
150
200
x (m)
0
50
100
150
200
250
300
y (m)
0
50
100
150
200
x (m)
0
50
100
150
200
250
300


y (m)
0
50
100
150
200
x (m)
0
50
100
150
200
250
300
y (m)
0
50
100
150
200
x (m)
0
50
100
150
200
250
300

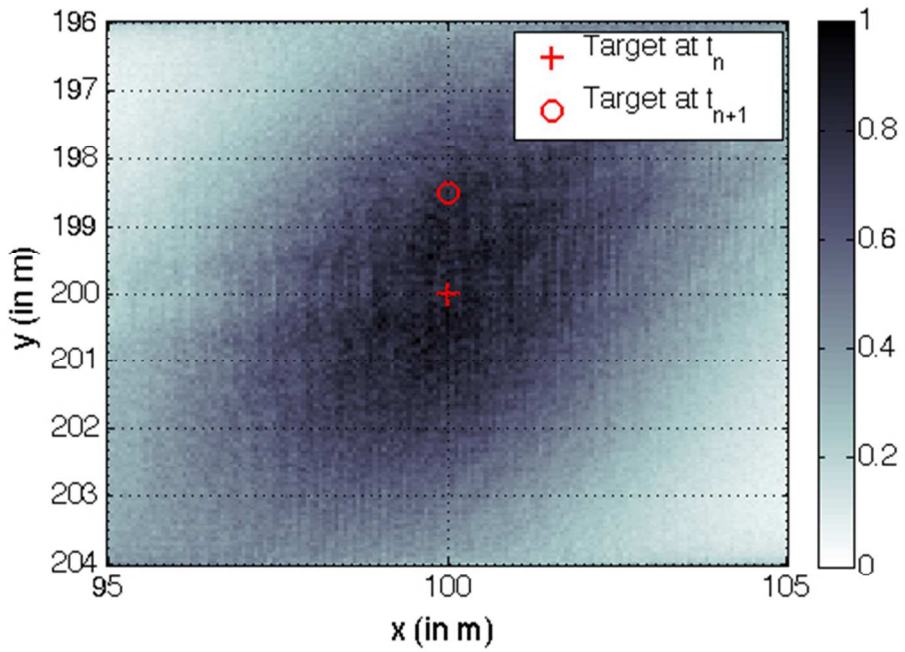

HISP static
HISP fish
HISP boat
PHD static
PHD fish
PHD boat
OSPA distance
time (s)
0
10
20
30
40
50
0
50
100
150
200
250


HISP static
HISP fish
HISP boat
PHD static
PHD fish
PHD boat
OSPA distance
time (s)
0
10
20
30
40
50
0
50
100
150
200
250


Transmission of the initial pulse E_0
Echo recording (Eq.(22))
Static component removal (Eq.(25))
Transmission of the unfocused time reversed echoes (Eq.(26))



Normalised energy of the back-propagated acoustical field using (a) traditional Time Reversal (Eq.~\eqref{TR}), and (b) unfocused Time Reversal (Eq.~\eqref{TRunfocus}).



Normalised energy of the back-propagated acoustical field using (a) traditional Time Reversal (Eq.~\eqref{TR}), and (b) unfocused Time Reversal (Eq.~\eqref{TRunfocus}).